Wealth Inequality, Wealth Constraints and Economic Performance*

Pranab Bardhan†
Samuel Bowles††
Herbert Gintis††
University of California, Berkeley†
University of Massachusetts, Amherst††

August 19, 1998

1 Introduction

If costlessly enforceable contracts regulate all the actions of economic actors that affect the well being of others, competitive equilibria are Pareto efficient regardless of the distribution of wealth. However, where actions such as risk taking and hard work are not subject to such contracts, the assignment of residual claimancy over income streams and control over assets—that is, the distribution of property rights—will affect the feasibility, cost, and effectiveness of contractual provisions and other incentive devices that may be used to attenuate the incentive problems arising from contractual incompleteness. In this situation, as we will presently see, some distributions of assets support efficient or near-efficient competitive allocations, while others preclude efficient contractual arrangements. The widespread use of tenancy contracts governing residential and agricultural property is an example of the latter. Thus where contracts are incomplete or unenforceable, as is the case in labor and credit markets, the distribution of wealth matters for allocative efficiency.

Contractual incompleteness and unenforceability arise when actors have information that is either private (others do not have it) or is inadmissi-

*We would like to thank P. Jeffrey Allen, Jean Marie Baland, Roland Bénabou, Timothy Besley, François Bourguignon, Michael Carter, Gregory Dow, Ernst Fehr, Karla Hoff, Eric Verhoogen and Elizabeth Wood for their helpful input in preparing this paper. We would also like to thank Jeffrey Carpenter, Jeff Dayton-Johnson, and Yongmei Zhou for research assistance and the MacArthur Foundation for financial support.

ble in judicial proceedings and hence cannot be used to enforce contracts. Contractual incompleteness may also arise where appropriate judicial institutions are lacking, as in the case of sovereign debt among nations, or where potential users of commons-type resources cannot easily be excluded from access. In these cases the distribution of wealth may affect allocative efficiency by its impact on

- residual claimancy over income streams and hence incentives for both an agent's own actions and the agent's monitoring of the actions of others;
- exit options in bargaining situations;
- the relative capacities of actors to exploit common resources;
- the capacity to punish those who deviate from cooperative solutions; and
- the pattern of both risk aversion and the subjective cost of capital in the population.

In this chapter we examine recent economic thinking about wealth effects on allocative efficiency in cases where information asymmetry, nonverifiability, or nonexcludablity of users, makes complete contracting infeasible.

It may be thought that the class of cases we are dealing with is not extensive, once the institutional setting is extended from one of competitive behavior governed by parametric prices to a more general environment in which private bargaining among small numbers of actors is feasible. Where property in assets may be readily traded and there are no impediments to efficient bargaining, inefficient assignments of control and residual claimancy rights over assets will be eliminated by voluntary exchange. This Coasean insight motivates the expectation that in competitive market economies, assets will be held by those who can use them most effectively, irrespective of their wealth.[1] If a tenant, for example, could make better use of the land as owner, the land should be worth more to the tenant than to the landlord, and hence one might expect the tenant to buy the asset.

But the very informational asymmetries that make some assignments of property rights more efficient than others also systematically impede the

[1] Grossman and Hart (1986):694 use this reasoning, for instance, to "explain asset ownership" while Hölmstrom and Tirole (1988) write that "contractual designs... are created to minimize transactions costs... This follows Coase's original hypothesis that institutions... can best be understood as optimal accommodations to contractual constraints..."

productivity enhancing reassignment of property rights. In particular, nonwealthy agents may be credit constrained, and hence may not find it possible to acquire those assets for which their exercise of residual claimancy and control rights would allow efficiency gains.

Hence the Coasean reasoning may not apply. Competitively determined property rights assignments may be technically inefficient in the standard sense that there exist alternative allocations that produce the same outputs with less of at least one input and not more of any. These inefficiencies may be attenuated by a nonmarket transfer of assets from wealthier to less wealthy actors.[2]

However, potentially offsetting efficiency losses may result from egalitarian asset transfers where, as will generally be the case, they result in a transfer of control over productive risk taking from less to more risk averse agents. An important productivity enhancing aspect of high levels of wealth inequality is that assets are controlled by agents who are close to risk neutral, and who thus choose a more nearly socially optimal level of risk.[3]

Wealth concentrations may support productivity enhancing allocations in other ways as well, as for example, in attenuating free rider problems in the monitoring of corporate managers by owners (Demsetz and Lehn 1985) or in more standard collective action problems (Olson 1965). Similarly, the ethic of egalitarian sharing that pervades many simple societies may reduce incentives for individual investment, as the returns, should they materialize, will be shared while the costs will be individually borne (Hayami and Platteau 1997). There can thus be no *a priori* conclusion concerning the efficiency effects of egalitarian asset redistribution.

The reader will note that in our argument, we have used the term 'productivity enhancing' in place of the more familiar term 'Pareto efficient.' *We define a policy as productivity enhancing if the gainers could compensate the*

[2]Several recent studies suggest that some egalitarian redistributions can have positive efficiency effects. See the works cited above as well as Legros and Newman (1997), Moene (1992), Manning (1992), Mookherjee (1997), and Bénabou (1997). The underlying argument has many precursors including the early nutrition-based efficiency wage theories.

[3]We shall assume in this paper that the 'socially optimal risk level' for individual projects is that which maximizes expected return. This is strictly true only if there are many individual projects and the returns to individual projects are uncorrelated, or if projects are correlated but there is some macroeconomic mechanism for smoothing returns across periods at zero cost.

Suboptimal risk choices in this sense are socially inefficient, but perhaps more important, such suboptimal risk choices can lead to dynamic inefficiencies if risk taking promotes the emergence and diffusion of new products and more advanced technologies. We do not model this effect here.

losers and still remain better off, except that the implied compensation need not be implementable under the informational conditions and other incentive constraints in the economy. We will use this term for the following reasons.

Since we are analyzing distributions of assets (and hence of access to income and well being) the usual standard of Pareto efficiency will generally be inapplicable. A mandated transfer of an asset to an erstwhile employee that results in technical efficiency gains by reducing monitoring inputs, for example, is unlikely to represent a Pareto improvement without compensation from the gainer to the loser, and the required compensation, if implemented, would reverse the effect of the initial asset transfer and dampen the associated incentive effects. Indeed, were it a Pareto improvement, the transfer would be readily accomplished through private exchange, as long as bargaining over the assignment of property rights is unimpeded.

If the Pareto criterion is too stringent, the cardinalist alternative, based on 'aggregate utility' is insufficiently so, for it may count an egalitarian redistribution as efficiency enhancing simply by dint of the diminishing marginal utility of income, even if the redistribution results in technical efficiency losses.[4] Transferring income from the rich to the poor could thus readily pass an efficiency test even if it were accomplished using very 'leaky buckets' (to use Okun's expression for inefficient transfer mechanism). By contrast, commonly used net output measures avoid reference to individual utility altogether and thus preclude evaluation of welfare-relevant variations in work effort and risk, a serious shortcoming for an analysis in which these non-contractible but welfare-relevant behaviors play a central role. Similarly, the widely used joint surplus maximization criterion is applicable only where utilities are assumed to be linearly additive, requiring risk neutrality, and thereby failing to address the central questions concerning risk behavior and insurance.

The technical efficiency criterion—more output with less of at least one scarce input and not more or any—is uncontroversial if inputs are defined sufficiently broadly, but fails to give a decisive ranking in the case of most redistributions.

Finally, the usual 'compensation' criteria, which consider a change effi-

[4] Bénabou (1996) addresses this problem by developing a measure of "pure economic efficiency which fully incorporates investment effects, labor supply effects and insurance effects but does not involve any interpersonal comparisons of utility." (p. 13). He distinguishes between the intertemporal elasticity of substitution of individuals, which is incorporated in his measure, along with risk aversion, and society's possibly egalitarian evaluation of the appropriate interpersonal elasticity of substitution, which he treats separately as a normative, but not pure efficiency related, measure.

cient if the gainers could compensate the losers, are often inapplicable, since the improved incentives that account for the efficiency enhancing properties of the redistribution would be lost if compensation were actually made. For instance, if transferring land to a landless peasant is viable because it improves the wealth of the peasant and thus reduces the subjective costs of increased risk to which the peasant as landowner is subject, then obliging the peasant to compensate the former landlord may eliminate the effect of the transfer in reducing risk aversion, leaving the peasant worse off than prior to the transfer.

For these reasons we prefer the productivity enhancement criterion over the more traditional alternatives. Of course many productivity enhancing redistributions also satisfy the technical efficiency, joint surplus maximization, Pareto-efficiency, and other conventional criteria.

In Section 2 we take up the first of our major cases: credit market misallocations when some agents are risk neutral but wealth constrained. Contractual incompleteness here arises from asymmetric information concerning risk taking behavior, and unenforceability arises from limited liability restrictions. The main result is that because inferior projects will be funded and superior projects not implemented when some agents have limited wealth, a redistribution of assets may be productivity enhancing in the above sense. In Section 3 we consider the persistence of inefficient contractual relationships governing land (including insecurity of tenure, sharecropping, wage labor and others). The main result of this section is that given the incompleteness of credit markets and other aspects of agrarian social structures, the market assignment of residual claimancy and control rights will often not be to the parties that can make the best use of the land.

In Section 4 we take up the consequences of redistributive policies for risk taking and risk exposure when nonwealthy agents are risk averse. We identify both positive and negative effects. Increasing the wealth of the nonwealthy will support higher levels of risk taking among this group, but concentrating the wealth of this group in ownership of the assets with which they work will likely have the opposite effect. And a redistribution of residual claimancy and control from the wealthy to less wealthy individuals is likely to induce a reduction in the aggregate level of risk taking, with likely adverse consequences for innovation and efficiency in the long run. We consider the role of tax-subsidy policies and insurance against exogenous risk in attenuating these adverse efficiency consequences.

Section 5 extends the analysis of Section 2 to the problem of team production. We explore the allocative distortions that may arise when, for example, employee effort is not contractible because the relevant information is nonverifiable and production team members are wealth constrained. We investigate the allocative implications of a reassignment of residual claimancy and control rights to employees, along with the introduction of mutual monitoring. We show that an asset transfer to team members may be productivity enhancing even when teams are large.

Our final case, addressed in Section 6, addresses the relationship between wealth distributions and the provision of local public goods. We explore possible relationships between inequalities in initial endowments and the ability of a group of individuals to solve collective action problems on the local commons, where allocative inefficiencies may arises from problems of nonexcludability.

There are a number of ways that wealth inequalities may depress productivity that are not addressed in what follows. Perhaps most important is the possibility that high levels of inequality (of wealth or income) may induce political instability and insecurity of property rights, which in turn depress investment and productivity growth.[5] The fact that the poor suffer productivity-reducing nutritional and other health problems that might be attenuated by an egalitarian asset redistribution (Leibenstein, 1957; Dasgupta and Ray, 1986,1987) is likewise not addressed.

We do not address processes of saving, human capital investment, bequests, or the political processes and policies that influence the time path of wealth redistributions, as these are dealt with elsewhere in this volume (see especially Piketty). Thus while we take account of the fact that the feasibility of an asset redistribution requires that it be supportable in a competitive equilibrium, we do not study the long term evolution of asset distributions under the influence of the forms of contractual incompleteness and public insurance and redistribution policies that we consider.[6] Finally, while we believe that the noncontractible effort, risk taking and other actions on which our argument hinges are critical to sustaining high levels of economic performance and that aligning incentives so that actors are residual claimants on the consequences of their actions may have a substantial

[5]For the first relationship see Alesina and Perotti (1996), Barro (1996), Keefer and Knack (1995), Perotti (1992) and Perotti (1993) and for the second all of the above plus Svensson (1993) and Venieris and Gupta (1986). However, one's concern that these cross-sectional results may provide little insight regarding relationships operating over time within countries is suggested by the failure of a measure of inequality to have the predicted negative coefficient in a country fixed effects model predicting investment (Benhabib and Speigel 1997). In fact their coefficient is positive and significant.

[6]Robinson (1996) surveys the interactions between wealth distributions and political equilibria affecting macroeconomic policies, explicit redistribution, policies toward labor unions and the distribution of political rights.

impact on productivity, we do not address the size of the relevant effects.[7]

2 Wealth and Efficiency when Risk is Non-Contractible

This section uses a simple model to illustrate the result that nonwealthy agents are disadvantaged in gaining access to credit in that their projects may go unfunded even when less socially productive projects of wealthy producers are funded.

A series of recent papers following the early work of Loury (1981) have analyzed the credit limitations faced by the nonwealthy and their efficiency effects.[8] These models show that when it is impossible to write complete state-contingent contracts, the equity a producer commits to a project reduces adverse selection and moral hazard problems by signaling the project's quality and increasing the producer's incentive to work hard and take the appropriate level of risk. However if there is a wealth constraint that limits agents' equity to their wealth holding, the nonwealthy may not be able to signal the quality of their projects and commit themselves to taking appropriate levels of effort and risk. Under these conditions, a redistribution of wealth from the wealthy to the nonwealthy may improve the efficiency of the economy fostering the substitution of more efficient production by the nonwealthy for less efficient production by the wealthy.

Are such credit constraints in fact operative? Several studies have shown that low-wealth producers in developing countries may be entirely shut out of credit markets or out of labor or land rental contracts that elicit high effort (Laffont and Matoussi 1995, Carter and Mesbah 1993, Barham, Boucher and Carter 1996, Carter and Barham 1996, Sial and Carter 1996). Other studies in low-income countries (Rosenzweig and Wolpin 1993) show that net worth strongly affects farm investment, and low wealth entails lower return to independent agricultural production (Rosenzweig and Binswanger 1993, Laffont and Matoussi 1995). Similarly, low net worth appears to depress labor market opportunities (Bardhan 1984).

[7] We refer to the relevant empirical literature throughout. Lazear (1996) studied a shift from hourly to piece rates in a large U. S. company and found "extremely large" productivity effects. Similarly a study of the effects of changing from salaried management to management by a residual claimant (as well as changes in the other direction) in a large chain of U. S. restaurants revealed strong residual claimancy effects (Shelton 1957).

[8] See Stiglitz (1974), Gintis (1989), Stiglitz (1989), Banerjee and Newman (1993), Rosenzweig and Wolpin (1993), Galor and Zeira (1993), Bowles and Gintis (1994), Barham, Boadway, Marchand and Pestieau (1995), Hoff and Lyon (1995), Hoff (1996b), Legros and Newman (1996), Aghion and Bolton (1997), Bénabou (1996), Piketty (1997).

Turning to the advanced economies, Blanchflower and Oswald (forthcoming) found that an inheritance of $10,000 doubles a typical British youth's likelihood of setting up in business, and another British study (Holtz-Eakin, Joulfaian and Rosen 1994b, Holtz-Eakin, Joulfaian and Rosen 1994a) found an elasticity of self-employment with respect to inherited assets of 0.52, and that inheritance leads the self-employed to increase the scale of their operations considerably. A third British study (Black, de Meza and Jeffreys 1996) found that a 10% rise in value of collateralizable housing assets in the United Kingdom increases the number of startup businesses by 5%. Evans and Jovanovic (1989) find that among white males in the U. S., wealth levels are a barrier to becoming entrepreneurs, and that credit constraints typically limit those starting new businesses to capitalization of not more than 1.5 times their initial assets: "most individuals who enter self-employment face a binding liquidity constraint and as a result us a suboptimal amount of capital to start up their businesses." (810).[9]

Consistent with the hypothesis that the poor are credit constrained is the strong inverse relationship between individual incomes and rates of time preference. Hausman (1979) estimated rates of time preference from (U.S.) individual buyers' implicit tradeoffs between initial outlay and subsequent operating costs in a range of models of air conditioners. He found that while high income buyers exhibited implicit rates of time preference in the neighborhood of the prime rate, buyers below the median income level exhibited rates five times this rate. Green, Myerson, Lichtman, Rosen and Fry (1996) elicited (hyperbolic) discount rates from high and low income respondents in the U. S. using a questionnaire method. The low income group's estimated rates were four times the high income group. In both the Green et al. and the Hausman study the elasticity of the rate of time preference with respect to income was approximately minus one.

We present a simple model illustrating this phenomenon.

Consider a set of 'producers,' each of whom has access to an investment project the returns to which depend on the level of risk assumed by the producer. We assume the projects themselves cannot be exchanged among agents. Producers must therefore finance the project out of their own wealth or by borrowing. We assume all agents are risk neutral, and credit markets are competitive, in the sense that in equilibrium lenders receive an expected return equal to the risk free interest rate.

The results below are true if three conditions hold: (i) the level of risk assumed by a producer is private information and hence cannot be contracted

[9] See also Evans and Leighton (1989).

form by a lender; (ii) any loan contract has a limited liability provision so that the promise to repay a loan may be unenforceable; and (iii) there a minimum project size.

In this situation we show the following:

- socially productive projects of low wealth producers may not be funded and hence may not be undertaken;
- Wealthy agents' relatively less productive projects may be funded in circumstances where less wealthy agents' relatively more productive projects are not funded;
- Wealthy agents will fund larger projects than less wealthy agents;
- If some producers are credit-constrained, a redistribution of wealth from lenders to such producers will be productivity enhancing; conversely if some but not all producers are size constrained, an asset transfer from a size constrained producer to a wealthier unconstrained producer may be productivity enhancing;
- If producers have projects of differing quality, there may exist an productivity enhancing asset transfer from a wealthy producer with a low-quality but profitable project to a credit-constrained producer with a high-quality project;
- If there are decreasing returns to scale, and if some but not all producers who are credit-constrained, there exist productivity enhancing redistributions from wealthy to nonwealthy agents;

Consider a project for which the relationship between risk and expected return is $\phi(p)$, where $p \in [0,1]$ is the probability of failure, a measure of the riskiness of the project, and $\phi(p)$ is the expected return net of all costs except capital costs of a project of unit size and quality. The riskiness of the project depends on the choice of technique (the type of seed planted, or the speed of operation of equipment) and level of effort or care taken by the producer. We assume $\phi(p)$ is inverted u-shaped, meaning $\phi'' < 0$ and $\phi'(p^*) = 0$ for some $p^* \in (0,1)$.[10]

[10] We take this to be the most plausible shape for the following reasons. First, production techniques that offer positive expected return are likely to involve a strictly positive level of risk. Hence expected return is an increasing function of risk for low levels of risk. However firms usually have access to production techniques that have very high returns when successful, but with a low probability of success (e.g. a firm may lower costs by

Suppose producer with wealth $\mathbf{w}$ whose project of size $\alpha \geq 1$ requires capital $\alpha\mathbf{k} > 0$, which is fully depreciated in one period, and has expected return $\alpha\beta\phi(p)$, where $\beta > 0$ is a parameter representing the *quality* of the project. Expected returns $\alpha\beta\phi(p)$ are shown in the top panel of Figure 1. Note that projects have minimum size of unity, but can be expanded with constant returns to scale above this minimum size.

Clearly p^* is the Pareto optimal risk level, since this maximizes the expected return to the project, and both producer and lender are risk neutral. We say the project is *productive* if

$$\beta\phi(p^*)/\mathbf{k} > 1 + \rho, \tag{1}$$

where $\rho > 0$ is the risk free interest rate. Thus a project is productive when the expected return per dollar of investment at p^* exceeds the return to a risk free security, and hence would be attractive to a risk neutral investor.

An *equity-backed loan* with equity c and interest rate r is a contract in which a producer with wealth $\mathbf{w}$ contributes equity $c \leq \mathbf{w}$ towards financing the project and the lender supplies the producer with the remainder $\alpha\mathbf{k} - c$. The producer then repays $(1 + r)(\alpha\mathbf{k} - c)$ if the project is successful, and nothing otherwise. We assume the credit market is competitive and there is a perfectly elastic supply of risk neutral lenders at the risk free interest rate ρ. A producer can thus always obtain an equity-backed loan so long as the expected return to the lender is ρ.

We assume potential lenders know the producers's unit expected return schedule $\beta\phi(p)$, and can contract for particular levels of c and α, but the risk level p chosen by the producer is not subject to a costlessly enforceable contract, since the choice of technology and the effort level of the producer are both the private information of the producer. We thus have a principal-agent relationship in which the producer is the agent and the lender, who is the principal, knows that the interest rate r affects the agent's choice of non-contractible risk p. Given c, β and α, the producer chooses p to maximize the expected return minus the cost to the producer of financing the project, or expected profits which, recalling that nothing is repaid if the project fails, is

$$v = \alpha\beta\phi(p) - [(1 - p)(\alpha\mathbf{k} - c)(1 + r) + c(1 + \rho)], \tag{2}$$

not diversifying its product line, or by assuming the availability of particular production inputs). Such high-risk projects, which have low expected return, may be attractive to producers since lenders bear part of the losses in case of failure. Moreover, if producer effort reduces risk at a constant rate, and the cost of effort is convex, an inverted-u-shaped ϕ function may result independently of the choice of technology.

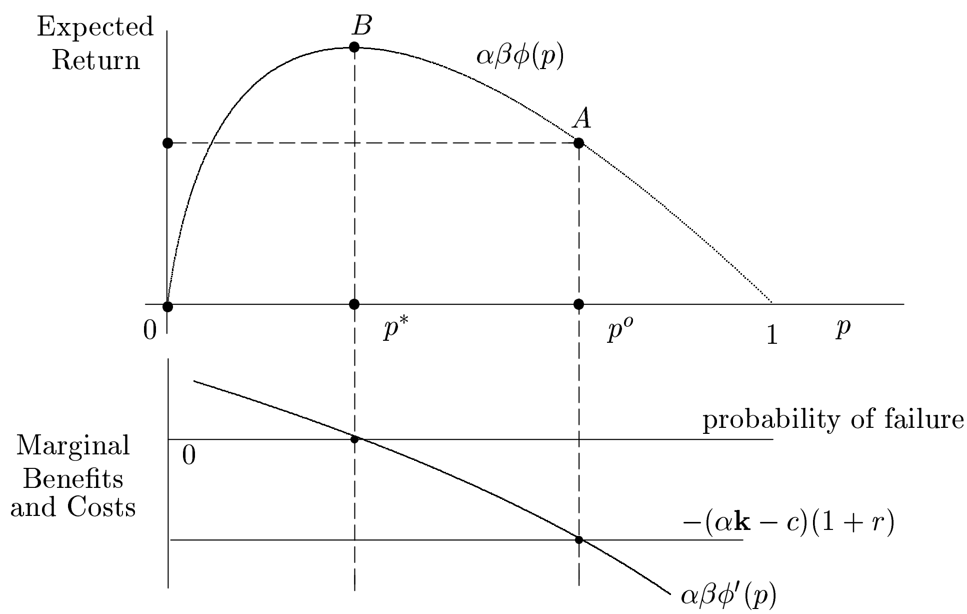

Figure 1: Excessive Risk in the Credit Market The equilibrium risk level is p^*.

for which the first order condition is

$$v_p = \alpha\beta\phi'(p) + (\alpha\mathbf{k} - c)(1 + r) = 0. \tag{3}$$

The result is illustrated in the lower panel of Figure 1, which reproduces equation (3). Note that the producer chooses p so that the marginal benefit from increased risk (namely the benefits of increase likelihood of non-repayment of the loan) or $(\alpha\mathbf{k} - c)(1 + r)$, is equated to the marginal cost of increased risk $\alpha\beta\phi'$.

It is clear from Figure 1 that if there is a positive level of borrowing, the chosen risk level p^o is excessively risky compared p^*, which is Pareto optimal, given that both borrower and lender are risk neutral.

The producer's best response function $p^o = p^o(r)$, which is the solution to (3) for various values of r, is depicted in Figure 2. This function is the locus of vertical tangent points on the producer iso-return schedules given by equating (2) to various levels of v. The producer iso-return loci are, by (3), is positively sloped for $p < p^o(r)$ and negatively sloped for $p > p^o(r)$, as shown at points A through C.[11]

[11] This is true because totally differentiating the producer iso-profit equation

$$\alpha\beta\phi(p^o) - [(1 - p^o)(\alpha\mathbf{k} - c)(1 + r) + c(1 + \rho)] = \underline{v},$$

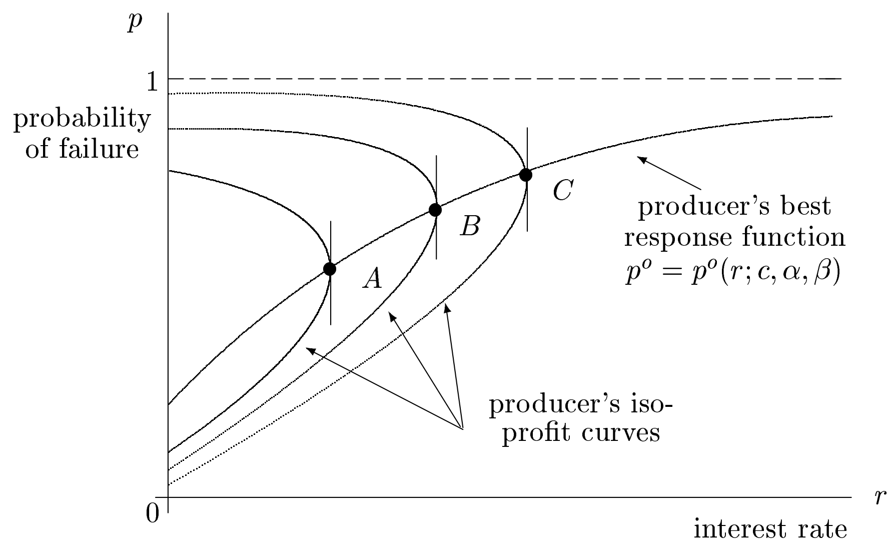

Figure 2: The Producer's Best Response Function: Risk Taken as a function of the Interest Rate.

Theorem 1. *The producer's best response function $p^o = p^o(r; c, \alpha, \beta)$ is an increasing function of r, shifts upward with an increase in project size α, and shifts downward with an increase in equity c or project quality β.*[12]

The intuition behind Theorem 1 is clear from Figure 1. An increase in r raises the marginal benefit of risk taking and induces a higher level of risk, as does a larger project (for a given level of equity, c) for it shifts a larger fraction of the cost of failure to the lender. Conversely an increase in the quality of the project increases the marginal cost of risk taking and induces less risk.

Figure 3 illustrates Theorem 1 in a particularly simple, but not implausible, case. Suppose the gross return to a successful project with probability p of failure is just $\alpha\beta p$, so $\phi(p) = p(1-p)$. Then the first order condition (3) becomes

$$\alpha\beta(1 - 2p) + (\alpha\mathbf{k} - c)(1 + r) = 0,$$

we get

$$\left.\frac{dp^o}{dr}\right|_{v=\underline{v}} = \frac{(1 - p^o)(\alpha\mathbf{k} - c)}{\alpha\beta\phi'(p^o) + (\alpha\mathbf{k} - c)(1 + r)},$$

the denominator of which is zero for p^o, by (3).

[12] The proofs of this and subsequent theorems are not included in this paper, but are available from the authors.

so rearranging the first order condition, the best response schedule is a linear function of r:

$$p^o = \frac{1}{2} + \frac{\alpha\mathbf{k} - c}{2\alpha\beta}(1+r).$$

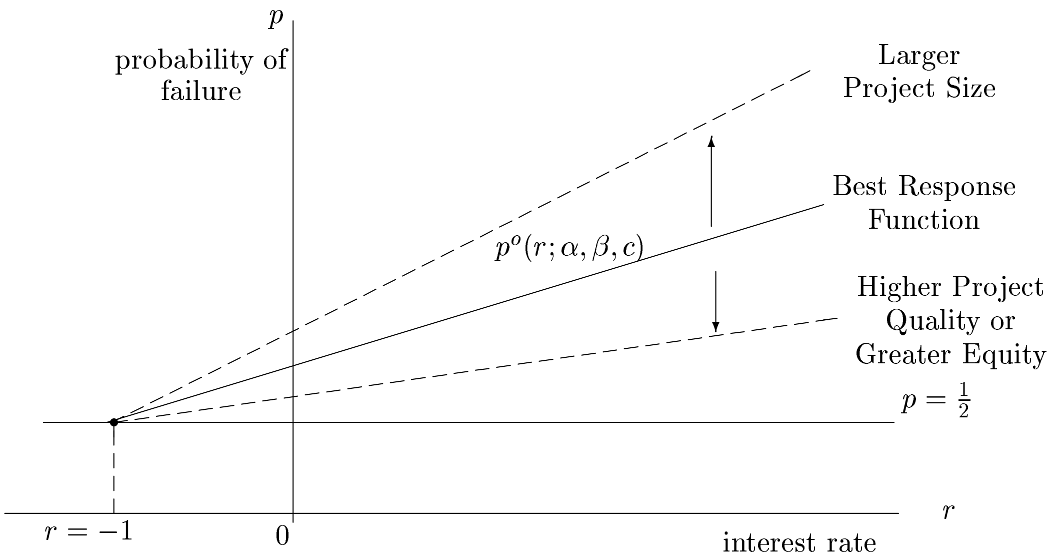

Figure 3: An Example: The Reaction of the Producer's Best Response Function to Changes in Project Size, Project Quality, and Producer Collateral when $\phi(p) = p(1-p)$. The Pareto-efficient risk level if $p = 1/2$.

It is clear from this expression that an increase in project size α increases both the intercept and the slope of this linear function, an increase in project quality β shifts down both the intercept and the slope, and an increase in equity c shifts down the intercept and slope.

How is r determined? Because the credit market is competitive, any equilibrium interest rate, failure rate pair $(r, p^o(r))$ must lie on the lender iso-expected-return schedule

$$(1+r)(1-p^o(r)) = 1+\rho, \tag{4}$$

which is an hyperbola, increasing and convex to the p-axis. The equilibrium r in p-r space must be consistent with the producer's best response function. The equilibrium contract will thus be the intersection of the producer's best response function $p^o(r)$ and the lenders iso-expected-return schedule at expected return ρ (should such an intersection exist). Thus, using (4), we have

$$r = \frac{p^o(r) + \rho}{1 - p^o(r)}, \tag{5}$$

Point B in Figure 4 illustrates such an equilibrium. This figure superimposes the lender's iso-expected return schedule on part of the Figure 2.

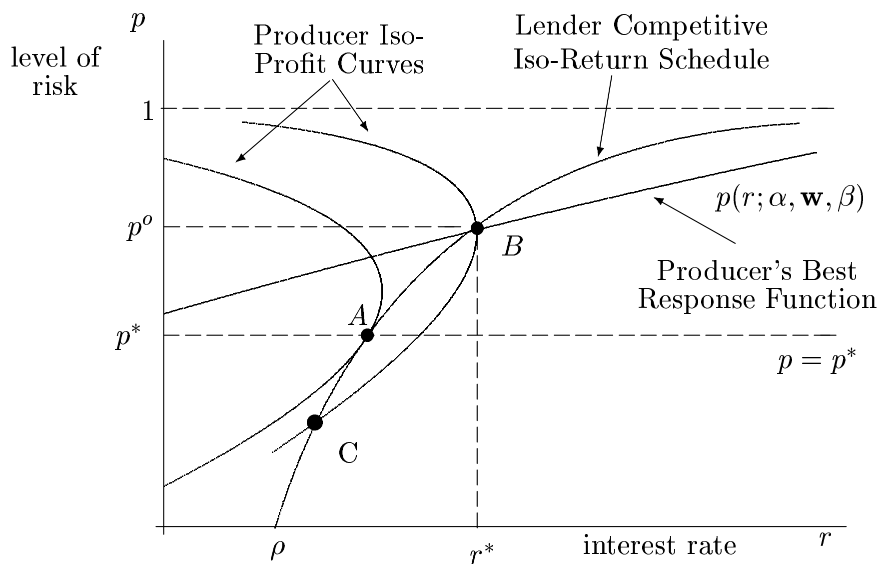

Figure 4: Equilibrium Choice of the Nominal Interest Rate r and Project Risk p: the equilibrium is a point B. The Pareto optimum satisfying the lender's participation constraint occurs at point A, as compared to the equilibrium at B in the figure.

We say an equity-backed loan is *feasible* if it offers a positive expected profit to the producer, and offers the lender an expected return at least as large as the risk free interest rate. The nominal interest rate r required to induce a lender to participate in the contract will be a function of equity c, project quality β, and project size α, since these parameters affect the producer's choice of (non-contractible) risk. As we have seen, the producer's best response function depends on both the project size α and the amount of the producer's own wealth c in the project, differing values of these variables entailing differing equilibrium levels of p and r as defined by equation (5). The producer thus varies α and c to maximize profits, taking account of the fact that his own best response function and hence the feasible interest rate r depends upon these variables.

The results presented below can be understood intuitively in terms of the following argument. First, since the producer is risk neutral and borrowing is more costly that self-financing (there are no incentive losses with self-financing), the producer always sets $c = \mathbf{w}$. Second, if the producer can choose the project size α without regard to the minimum size constraint $\alpha \geq 1$, there would be an optimal project size α^* and an optimal leverage ratio,

$$\lambda^* = 1 - \frac{\mathbf{w}}{\alpha^* \mathbf{k}},$$

namely one that balances the marginal benefits of borrowing (the ability to shift losses to the lender) against the marginal costs (the increased risk of failure occasioned by the producer's own choice of risk level under the incentive implied by the leverage ratio). Asset-poor producers, even if they are able to secure a loan, may be unable to implement λ^*, since for a poor producer it could be that $\alpha^* < 1$.

We say a project can be *fully financed* if a producer can obtain a loan to finance the project with no equity.

Theorem 2. *There are productive projects that cannot be fully financed.*

The reasoning behind Theorem 2 is that because the producer's best response is to select $p > p^*$, the fact that the project is productive, i.e., that $\beta\phi(p^*) > \mathbf{k}(1+\rho)$, does not insure that its expected return to the lender will equal or exceed $(1+\rho)$ when $p^o(r)$ rather than p^* is chosen by the producer. Thus the best response function of a non-wealthy producer with a productive project need not intersect the lender's competitive iso-return locus. Such producers will not be financed. There will thus exist productive projects that cannot be fully financed.

We are interested precisely in those cases where projects cannot be fully financed. In this situation, among producers with projects of equivalent quality, there will be three types. One group, whom we call the *credit-constrained*, will be unable to finance even a project of minimum size. A second, the *unconstrained*, have wealth sufficient to finance projects above the minimum size. They select a project size, given their wealth, to implement the optimal leverage ratio $\lambda^* = 1 - \mathbf{w}/\alpha^*\mathbf{k}$. A third group, the *size-constrained*, are able to finance a project only at the minimum size, and with $\lambda > \lambda^*$. They would indeed prefer a smaller project, as this would entail a lower, more nearly optimal, leverage ratio which in turn would induce a lower level of risk and a higher return on the project. But this is technologically precluded by the minimum project size $\alpha \geq 1$. We have

Theorem 3. *Suppose a producer with wealth* $\mathbf{w}$ *has a productive project, and selects risk level* p*, equity* $c \leq \mathbf{w}$*, and project size* $\alpha \geq 1$ *to maximize* *profits, subject to offering a lender an expected return equal to the risk free interest rate* ρ*. Suppose* r*,* c *and* α *are contractible but* p *is not, and the project cannot be fully financed. Then there are wealth levels* $0 < \underline{\mathbf{w}} \leq \overline{\mathbf{w}} < \mathbf{k}$ *such that the following assertions hold:*

(a) *for* $\mathbf{w} < \underline{\mathbf{w}}$*, an there is no feasible equity-backed loan, so the producer cannot finance the project.*

(b) *for* $\mathbf{w} \geq \underline{\mathbf{w}}$

- *(i)* *there is a feasible equity-backed loan, the producer contributes equity* $c = \mathbf{w}$*, and the project is undertaken;*
- *(ii)* *the probability of failure* p^o *chosen by the producer exceeds* p^**, so the choice of risk is not Pareto efficient;*

(c) *for* $\underline{\mathbf{w}} \leq \mathbf{w} < \overline{\mathbf{w}}$ *producers are size-constrained and*

- *(i)* *risk declines and project efficiency increases with increasing producer wealth;*
- *(ii)* *the rate of return on producer wealth,* $\alpha\beta\phi/\mathbf{w}$*, increases with increasing producer wealth;*

(d) *for* $\mathbf{w} \geq \overline{\mathbf{w}}$ *producers are unconstrained, the level of risk chosen and the rate of return is independent of wealth, and project size increases with producer wealth. Hence the total return on the project increases with producer wealth.*

Part (a) is just an extension of Theorem 2: if zero wealth producers are credit-rationed, the same will be true of those with any wealth level insufficient to yield a best response function that intersects the lender's iso-return locus.

The two wealth bounds may be interpreted as follows. The lower bound is the level of wealth for which the producer's best response function is tangent to the lender's competitive iso-return schedule. The upper bound is the wealth level for which the minimal size project implements the optimal degree of leverage, given the producer's wealth (i.e., $\overline{\mathbf{w}} = \mathbf{k}\lambda^*$).

To see why the lender contributes as much equity as possible to the project (part b(i) of the theorem), note that given the credit market competitive equilibrium condition (5), the producer's objective function can be written

$$v = \alpha[\beta\phi(p(\alpha, c)) - \mathbf{k}(1 + \rho)], \tag{6}$$

where $p(\alpha, c)$ is the risk level chosen by the producer who obtains an equity-backed loan with collateral c and project size α. So the expected financial cost of the project, including the debt service and the opportunity cost of supplying equity, is $\alpha\mathbf{k}(1+\rho)$ independent of how much of the producer's own wealth is invested. Since increasing c lowers p and $\phi'(p) < 0$ in the relevant region, the producer will set $c = \mathbf{w}$.

Part b(ii) holds because the lender pays some of the downside cost of the risk, and is clear from the producer's first order condition. Figure 4 illustrates this result. At the competitive equilibrium (point B) the producer's iso-profit curve has a vertical tangent, because it is a point on the producer's best response function–see Figure 2. However the lender's iso-return schedule has positive slope at this point. Since mutual tangency of the iso-return schedules of producer and lender is the condition for Pareto efficiency, B is not Pareto efficient. The Pareto efficient point corresponding to the lender's participation constraint is at point A in Figure 4, and always entail lower default probabilities and lower interest rates than the credit market equilibrium. The implied Pareto improvements are indicated by the lens-shaped region ABC. These points are infeasible, of course, as they are not on the producer's best response function. Note that the efficient contract locus is horizontal at level p^*, and has no points in common with the producer's best response function.

Part (c) of the theorem is true on the same reasoning motivating part (a). For size-constrained producers, investing more wealth does not change the expected financial cost of the project, but it does raise the expected returns because it induces a lower nominal interest rate and hence a lower level of risk. Over the specified range of wealth, lowering the leverage ratio reduces the allocational distortion identified in b(ii), as is clear from (3) and Figure 1.

The intuition behind part (d) is that $\mathbf{w} \geq \overline{\mathbf{w}}$ allows the producer to vary project size to implement the optimal leverage ratio, so increased wealth implies larger but otherwise identical projects.

Not surprisingly, wealthier producers may finance lower quality projects.

Theorem 4. *For any wealth* $\mathbf{w} > 0$*, there is a quality level* $\underline{\beta}(\mathbf{w})$ *such that the producer's project will be funded when* $\beta \geq \underline{\beta}(\mathbf{w})$*. The minimum wealth level* $\underline{\mathbf{w}}$ *at which a producer's project can be funded decreases as the quality* β *of the project increases. For a given producer wealth* $\mathbf{w}$*, optimal project size* $\alpha(\mathbf{w})$ *is an increasing function of project quality* β *for* $\beta \geq \underline{\beta}(\mathbf{w})$*.*

Theorems 3 and 4 are illustrated in Figure 5. The locus AA' separates

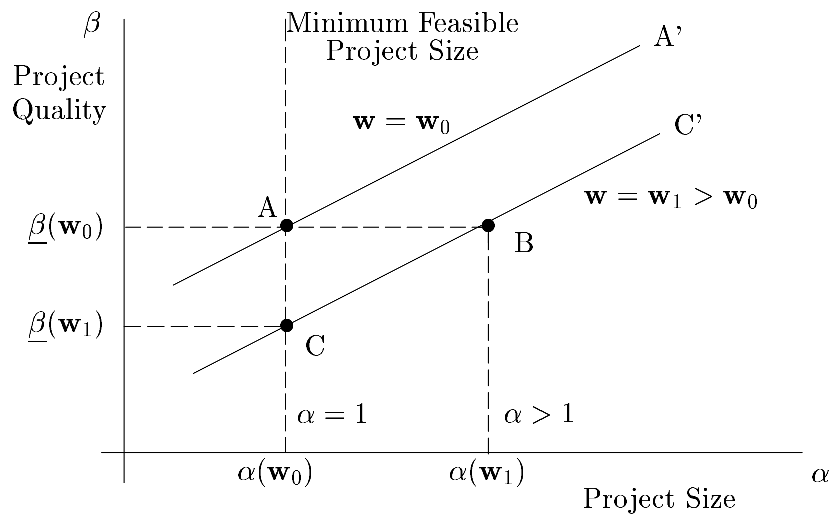

Figure 5: Greater wealth allows larger or lower quality projects to be financed

regions of feasible project (northwest and to the right of $\alpha = 1$) and infeasible project (southeast) for a producer of wealth $\mathbf{w}_0$, larger projects being fundable if they are superior quality. The locus CC' refers to a wealthier producer. It is clear that of two agents with projects of the same quality $\underline{\beta}(\mathbf{w}_0)$, the wealthier agent may have a larger project (of size $\alpha(\mathbf{w}_1)$ at point B) than the less wealthy (of size $\alpha(\mathbf{w}_0)$ at point A). Moreover, the wealthier agent can finance a lower quality project (of quality $\underline{\beta}(\mathbf{w}_1)$ at point C) than the minimum quality that can be financed by the less wealthy agent (which is of quality $\underline{\beta}(\mathbf{w}_0)$ at point A).

An implication of Theorem 3 is that the rate of return to wealth increases in the level of wealth, as is illustrated in Figure 6. Small wealth holders who are credit constrained receive ρ, while from (2) we see that the unconstrained receive an income of $v + c\rho$ giving them a rate of return of $v/\mathbf{w} + \rho$. Intermediate between these two cases, the size-constrained receive a return rising from the risk free rate to the unconstrained rate $\alpha\beta\phi(p(\lambda^*))/\mathbf{w} > 1 + \rho$, as wealth increases from $\underline{\mathbf{w}}$ to $\overline{\mathbf{w}}$. Figure 6 also illustrates the fact that the interest rate stipulated in the loan contract declines over the range $(\underline{\mathbf{w}}, \overline{\mathbf{w}})$ and is constant for higher levels of wealth. Because expected rates of return are lower for the less wealthy, wealth inequalities may grow over time. Notice, however, that the technological non-convexity constituted by the minimum project size $\alpha \geq 1$ is essential for the result: as Galor and Zeira (1993) point out, wealth constraints in credit markets alone do not bear this implication.

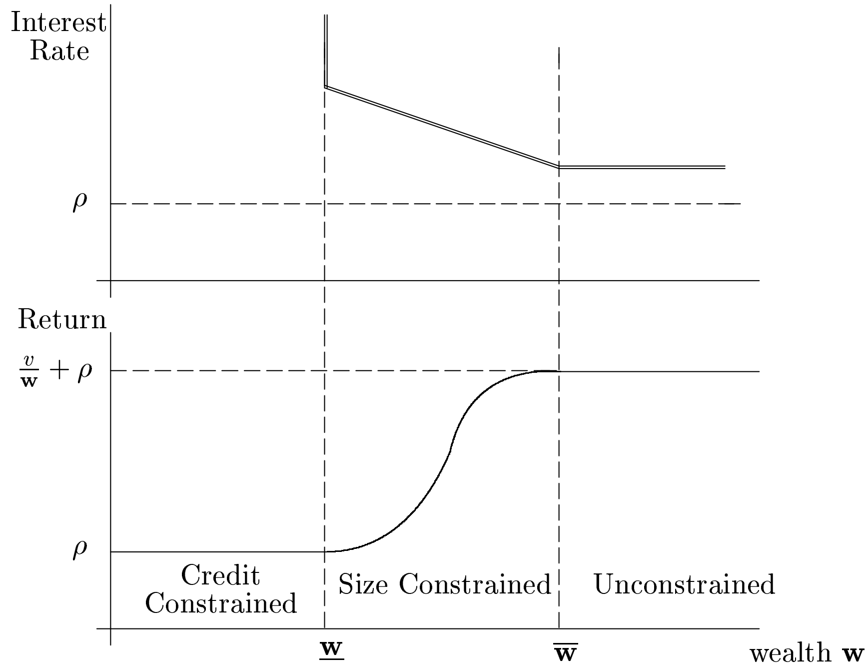

Figure 6: Rate of Return Increases with Producer Wealth

If projects of any size were possible, those with limited wealth would have small projects with leverage ratios λ^*, earning the same expected rate of return as the wealthy.

Because we treat the wealth levels of individuals as exogenous and so do not take account of savings incentives and the long run evolution of the wealth distribution we will mention only in passing an important implication of Theorem 3. Individuals holding wealth over (or near) the size constrained range $(\underline{\mathbf{w}}, \overline{\mathbf{w}})$ have enhanced savings incentives, and may even benefit more from saving than do the wealthy (despite the fact that the average returns on wealth rise with wealth, as we have seen). The reason is that for wealthholders in this range increased wealth supports more nearly optimal contracts, and allows higher returns on wealth, thus providing additional savings incentives. Thus if the expected return to holding wealth $\mathbf{w}$ is $r(\mathbf{w})$ the incentive to add one unit to one's wealth is $r(\mathbf{w}) + r'(\mathbf{w})\mathbf{w}$, where the second term may be sufficiently large that savings incentives for those in the size constrained range exceed those in the unconstrained range. If this is the case those in the size constrained range might over time save themselves out of their wealth constraints. Similar, though attenuated incentives exist for those whose low wealth leaves them credit constrained but for whom saving into the size constrained range is feasible. Thus we cannot say that wealth inequalities will grow overtime even when the return structure is as depicted in Figure 7, an adequate treatment of this requiring not only endogenous savings decisions (above), but an account of the risk return choices of individuals at various wealth levels, a topic to which we will turn shortly.[13]

Not surprisingly, asset redistribution may be productivity enhancing under the conditions given.

Theorem 5. *Suppose all projects are of the same quality. If some producers are credit-constrained, a redistribution of wealth from lenders to such producers will be productivity enhancing. If some producers are size constrained while others are unconstrained, an asset transfer from a size constrained producer to a wealthier unconstrained producer may be productivity enhancing.*

The first part of the theorem flows from the fact that lenders receive expected return $\alpha\mathbf{k}(1+\rho)$, on an investment of $\alpha\mathbf{k}$, while a credit-constrained producer would receive expected return $\alpha\beta\phi(p) > \alpha\mathbf{k}(1+\rho)$ on this investment. The second part of the theorem follows from the result that the unconstrained select less risky (and hence Pareto superior) levels of risk. Suppose the transfer is large enough to reduce the size constrained producer to the status of an excluded credit constrained non-producer. Then the effect of the transfer is simply to expand a project which will be operated in a less excessively risky manner, while eliminating a project being operated in more risky manner.[14]

We now relax two simplifying assumptions. Suppose, first, that project quality, β varies among producers. We know from Theorem 4 that where both project quality and wealth differ among producers, there may exist superior projects which are infeasible due to the insufficient wealth of the producer, while inferior projects are funded. This result motivates an important result, consistent with our initial claim that productivity enhancing egalitarian asset redistributions are possible.

[13] Zimmerman and Carter (1997) calibrate a dynamic portfolio choice model of this process using data from three regions of Burkina Faso. Birdsall, Pinckney and Sabot (1996) develop a model of high powered savings incentives for poor credit constrained households.

[14] Note that voluntary exchange will not bring about this inegalitarian productivity enhancing change, since the wealthier agent does not want to borrow, even at the risk-free interest rate.

Theorem 6. *Suppose individuals have projects of differing quality, and that some with productive projects are credit constrained while others are size constrained. Then there may exist an productivity enhancing asset transfer from a wealthy to a credit-constrained producer.*

The theorem follows trivially from the previous theorems. Under the stated conditions it will generally be possible to transfer wealth from an unconstrained producer with a relatively poor project (one with $\beta = \hat{\beta}$) to an initially credit constrained producer with a superior project (one with $\beta = \tilde{\beta} > \hat{\beta}$). To see how this might work, imagine that a wealthless producer has a high quality project. If the amount transferred is sufficient to give the previously credit-constrained producer an amount of wealth at least as great as $\overline{\mathbf{w}}$ without reducing the wealthy producer below this level, then the only effect of the transfer is to reduce the size of the inferior project by $\overline{\mathbf{w}}/\lambda$ while introducing a superior project of equivalent size. The resulting increase in aggregate output is $\overline{\mathbf{w}}(\tilde{\beta} - \hat{\beta})\alpha\phi(p(\lambda^*))$. The gains to the beneficiaries of the transfer obviously exceed the losses to the wealthy, and while it is not obvious how the losers could be compensated this is not required by our definition of a productivity enhancing asset transfer.

We now relax a second assumption. Suppose that each project faces a rising average cost schedule occasioned by the fact that while the minimum project size can be operated by the producer as residual claimant, larger project sizes require the employment of labor, occasioning a supervision cost. Specifically, suppose when the project size is $\alpha \geq 1$, the expected return schedule is

$$\alpha[1 - s(\alpha)]\beta\phi(p), \tag{7}$$

where $s(1) = 0$ where $s'(\alpha) > 0$ for $\alpha \geq 1$, and $s(\alpha) \geq 1$ for sufficiently large α. We have

Theorem 7. *Suppose there are decreasing returns to scale, given by the expected return schedule (7). Then if some producers are credit-constrained and other producers are sufficiently wealthy, there exist productivity enhancing redistributions from wealthy to nonwealthy agents.*

The intuition behind this theorem is as follows. For producer wealth $\mathbf{w}$ sufficient to secure an equity-backed loan, there is a profit maximizing project size $\alpha(\mathbf{w})$. Let $\mathbf{w}^*$ be the wealth level that maximizes the expected return per unit of capital $v(\alpha(\mathbf{w}), \mathbf{w})/\mathbf{k}\alpha(\mathbf{w})(1 + \rho)$. Then producers who have wealth $\mathbf{w} > \mathbf{w}^*$ can transfer $\mathbf{w} - \mathbf{w}^*$ to credit-constrained producers, which may allow them to operate at project size $\alpha(\mathbf{w}^*)$.

In sum, allocational distortions associated with loan contracts arise because limited liability protects borrowers from downside risks so that borrowers are not full residual claimants on the results of their actions. For the credit constrained the distortions are particularly great as their limited wealth induces them to assume a lower residual claimancy share than those with greater wealth. Those with sufficient wealth to assume full residual claimancy—by forgoing borrowing—adopt the Pareto efficient level of effort and risk choice. Where asset redistributions affect a redistribution of residual claimancy, as we have seen, they may attenuate or eliminate the above allocational distortions.

3 Land Contracts

Among the most important, and studied, cases of contractual incompleteness and wealth effects on allocative efficiency concern agricultural tenancy, whose harmful effects on productivity were famously lamented by not only de Tocqueville (in our headquote) but John Stuart Mill and Alfred Marshall as well. In many poor countries the empirical evidence suggests that economies of scale in farm production are insignificant (except in some plantation crops, and that too more in processing and marketing than in production) and when accompanied by appropriate insurance and credit institutions, the small family farm may often be the most productive unit of production.[15] Yet the violent and tortuous history of land reform in these countries suggests that there are numerous roadblocks on the way to a more efficient reallocation of land rights. Why do the large landlords not voluntarily sell their land to small family farmers and use their market and bargaining power to acquire much of the surplus arising from this reallocation?

First the small farmer, as a nonwealthy agent, faces the disadvantages in credit and insurance markets described in Sections 2 and 4, and hence is often not in a position to buy or profitably rent more land.[16] Second, land as an asset may serve some special functions for the rich that the poor

[15] See, for instance, Berry and Cline (1979) and Prosterman and Riedinger (1979), Ch. 2. For a more recent summary of the evidence and the methodological shortcomings of the empirical studies, see Binswanger, Deininger and Feder (1995). While there are many reasons for this regularity, among them are the agency problems we have stressed above (Shaban 1987, Laffont and Matoussi 1995).

[16] Mookherjee (1997), in a model with asymmetric information, thus generating potential informational rents for the supplier of unobserved effort, provides additional reasons why there will be no scope for mutually profitable land sales from landlords to tenants, or farm laborers.

are less capable of using and therefore are not reflected in the prices offered by the latter. For example, holding land may offer some tax advantages or speculative opportunities or a generally safe investment vehicle for the rich (particularly, when non-agricultural investment opportunities are limited or too risky) that are not particularly relevant for the small farmer. Similarly, large land holdings may give their owner special social status or political power in a lumpy way, so that the status effect from owning 100 hectares, for instance, is larger than the combined status effect accruing to 50 new buyers owning two hectares each. Binswanger et al. (1995) point out that land is often used as preferred collateral in the credit market and thus serves as more than just a productive asset. The asking price for land then may be above the capitalized value of the agricultural income stream for even the more productive small farmer, rendering mortgaged sales uncommon—since mortgaged land cannot be used as collateral to raise working capital for the buyer.

For all these reasons land ownership does not pass from the large to the small farmer and, accordingly, the land market is very thin. In rich countries a large part of land transactions may be related to the life-cycle: the elderly sell land to buyers at an accumulating stage in their life-cycle. In the more inter-generationally close-knit families in poor countries such life-cycle related land transactions are rare. More often in poor countries land sales go the way opposite to what is suggested by the evidence of the more efficient small farmer: land passes from distressed small farmers to landlords and money-lenders. This tendency increases as the traditional reciprocity-based risk-coping mechanisms weaken, and farmers may have to depend more on land sales in times of crisis.

We will analyze how the initial wealth distribution affects static efficiency in the tenancy market. We point to the possibility of a 'tenancy ladder' for tenants facing different wealth constraints. If risk neutral tenants are not wealth constrained, or if output does not depend strongly on tenant effort, the landlord will be able to devise a contract securing a Pareto efficient level of tenant effort, even where tenant effort is not verifiable. However

- where tenants have little wealth or where expected output depends strongly on tenant effort, inefficient contracts will obtain, with the resulting degree of allocative inefficiency varying inversely with the wealth of the tenant;
- for this reason a transfer of wealth to asset poor tenants may be productivity enhancing, even if the amount transferred is insufficient to permit the tenant to become the owner.

We use a principal-agent model that emphasizes, along with moral hazard, a limited liability constraint, i.e. the tenant is liable up to his own wealth level $\mathbf{w} \geq 0$.[17] We abstract from risk sharing issues and assume both the tenant and the landlord to be risk neutral (we address questions of risk subsequently).

Consider a plot of land that requires an input of one unit of labor and yields an output of either H or $L < H$. The probability of the H-output depends on the tenant's effort, and without loss of generality we may let this probability of the good state simply be $e \in [0, 1]$, where e is the effort applied to the land by the laborer. Suppose the owner of the plot, the landlord, hires a tenant to supply the labor. We assume that effort e has disutility $d(e)$ to the tenant, where $d(e)$ is increasing and convex, with $d(0) = d'(0) = 0$.

Let us start with the benchmark case where e is fully observable. Given the convexity of $d(e)$, there exists a unique value of e that maximizes the combined payoff of the landlord and the tenant, $eH + (1-e)L - d(e)$. Let us denote this first-best e as e^*. It is easy to see that the optimal level of effort will be that for which

$$d'(e^*) = H - L, \tag{8}$$

or the expected marginal product of effort equals the marginal disutility of effort. If e is fully observable, the landlord can offer the tenant a take-it-or-leave-it contract, which pays the latter's (given) reservation income $m \geq 0$ when e is observed to be e^*, and 0 if e is observed to be anything else.

But effort is not observable, so to motivate the tenant to supply an appropriate level of effort, the landlord offers the tenant a contract (h, l), under which the tenant pays the landlord $H - h$ when H is realized and $L - l$ when L is realized, the tenant retaining the remainder of the output, h in the good state and l in the bad. We assume the tenant has wealth $\mathbf{w} \geq 0$, and is subject to limited liability, so the constraints $h + \mathbf{w} \geq 0$ and $l + \mathbf{w} \geq 0$ cannot be violated. Only the latter inequality is of interest, however, since under any incentive scheme chosen by the landlord h will be greater than l (see below). Thus to induce effort the landlord must design a contract that satisfies an incentive compatibility constraint (ICC), a participation constraint (PC), and a limited liability constraint (LLC). That is, given such a contract, which cannot require a payment by the tenant larger than

[17] Early theoretical models in the literature with the limited liability constraint are those of Sappington (1983) and Shetty (1988). In a more recent paper Laffont and Matoussi (1995) use a data set from the region of El Oulja in Tunisia to support their theoretical result that financial constraints have a significant impact on the type of tenancy contract chosen, and hence on productive efficiency.

his wealth $\mathbf{w}$ (LCC), the tenant chooses e as a best response (ICC), and receives an expected utility no lower than m (PC). The landlord thus varies h and l to maximize

$$\pi = \max_{h,l} e(H-h) + (1-e)(L-l)$$

subject to

$$\begin{array}{ll} \text{(PC)} & eh + (1-e)l - d(e) \geq m, \\ \text{(ICC)} & e \in \arg\max_{\tilde{e}} \tilde{e}h + (1-\tilde{e})l - d(\tilde{e}), \\ \text{(LLC)} & l + \mathbf{w} \geq 0. \end{array}$$

The ICC must always hold as an equality, so it can be replaced by its corresponding first order condition, namely

$$h - l = d'(e). \tag{9}$$

This equality constraint can be used to eliminate the variable h, reducing the landlords maximization problem to

$$\pi = \max_{l,e} e(H - l - d'(e)) + (1-e)(L-l) \tag{10}$$

subject to

$$\begin{array}{llcl} \text{(PC)} & l + ed'(e) - d(e) - m & \geq & 0, \\ \text{(LLC)} & l + \mathbf{w} & \geq & 0. \end{array}$$

To solve this problem, we form the Lagrangian

$$\begin{aligned} \mathcal{L} = {} & e(H - l - d'(e)) + (1-e)(L-l) + \lambda(l + \mathbf{w}) \\ & + \mu(l + ed'(e) - d(e) - m), \end{aligned}$$

The first order conditions for this system are

$$H - L - d'(e) - (1-\mu)ed''(e) = 0 \tag{11}$$

$$\lambda + \mu = 1 \tag{12}$$

$$\lambda(l + \mathbf{w}) = 0 \tag{13}$$

$$\mu(l + ed'(e) - d(e) - m) = 0, \tag{14}$$

plus the two inequality constraints (PC) and (LLC). We will see that depending on the tentant's wealth level, one or both of these constraints will be binding.

Suppose first $\mu = 0$, so the PC is not binding. Then (11) determines an effort level e_* such that

$$H - L = d'(e_*) + e_* d''(e_*), \tag{15}$$

and since $\lambda = 1$, $l = -\mathbf{w}$, so the contract stipulates that in the bad state the tenant surrenders his entire wealth to the landlord. The PC must also not be violated, so

$$\mathbf{w} \leq \mathbf{w}_* \equiv e_* d'(e_*) - d(e_*) - m. \tag{16}$$

Thus for wealth in the range $[0, \mathbf{w}_*]$, the LCC constrains, the PC does not (the tenant enjoys a level of utility superior to his fallback position; i.e., he earns a rent, except at $\mathbf{w} = \mathbf{w}_*$), and effort is fixed at e_*. Finally the landlord's return is an increasing function of the tenant's wealth $\mathbf{w}$. This can be seen from (10), which by substituting $l = -\mathbf{w}$ now becomes

$$\pi = \mathbf{w} + L + e_*(H - L - d'(e_*)).$$

Comparing (8) and (15) we see that $d'(e_*) < d'(e^*)$, so the convexity of $d(e)$ implies $e_* < e^*$. We conclude that the low wealth tenant's effort level is less than the Pareto optimal level—that which maximizes the joint surplus.

Suppose, by contrast, that $\lambda = 0$, so the LLC is not binding. In this case $\mu = 1$, the PC holds as an equality (the tenant receives his fallback). Then (11) implies $H - L = d'(e)$, so the effort level is just e^*, as in (8), where effort is fully observable and the joint surplus is maximized. We see that because the PC is satisfied as an equality (using 14), the tenant's income in the low output state is now given by

$$l = m - e^* d'(e^*) + d(e^*),$$

and to satisfy the LLC we must have

$$\mathbf{w} \geq \mathbf{w}^* \equiv e^* d'(e^*) - d(e^*) - m. \tag{17}$$

For $\mathbf{w} \in [\mathbf{w}^*, \infty)$ effort is fixed at e^*, the low output state payment is fixed at $l = -\mathbf{w}^*$, and the LLC is not a binding constraint. Tenant wealth levels at least as great as $\mathbf{w}^*$ thus allow a contract supporting a Pareto-optimal outcome.

The only remaining possibility in (11)–(14) is that both λ and μ are nonzero. This case, in which both the LLC and the PC bind, represents a solution to the landlord's maximization problem only for $\mathbf{w} \in [\mathbf{w}_*, \mathbf{w}^*]$.[18] In this range, from the LLC we know that $l = -\mathbf{w}$ and from the PC we know that effort satisfies

$$\mathbf{w} = e_0 d'(e_0) - d(e_0) - m. \tag{18}$$

Notice here that $de_0/dw = 1/e_0 d''(e_0) > 0$, so $e_0(\mathbf{w})$ increases from e_* to e^* as $\mathbf{w}$ moves from $\mathbf{w}_*$ to $\mathbf{w}^*$. Also, l falls from $-\mathbf{w}_*$ to $-\mathbf{w}^*$ over this range, while h rises from $d(e_*) - \mathbf{w}$ to $d(e^*) - \mathbf{w}$.[19] The effect of the steepening payment gradient $h - l$ is to align tenant effort incentives more closely with the payoff to the landlord, approaching the Pareto optimal outcome as $\mathbf{w}$ approaches $\mathbf{w}^*$.

We thus have

Theorem 8. Tenancy Ladders and Inefficient Contracts *Where tenant effort e affects expected output and is not contractible, and where the tenant's wealth level $\mathbf{w}$ is the maximum liability to which the tenant may be exposed by the terms of the landlord's output-contingent rental contract, there exist wealth levels $\mathbf{w}_*$ and $\mathbf{w}^* > \mathbf{w}_*$ such that the landlord's optimal contract has the following properties:*

(i) for $\mathbf{w} \geq \mathbf{w}^*$ the tenant chooses a Pareto optimal effort level e^*, the tenant's expected return equals the tenant's reservation position, and the landlord secures the entire surplus;

(ii) for $\mathbf{w} < \mathbf{w}_*$ the tenant chooses effort $e_* < e^*$ independently of $\mathbf{w}$, the tenant's expected return is superior to the reservation position and declining in the tenant's wealth, while the landlord's share of output is increasing in the tentant's wealth; and finally,

(iii) for $\mathbf{w}_* \leq \mathbf{w} \leq \mathbf{w}^*$, tenant effort is rising in tenant wealth (from e_* to e^*), the tenant's expected return equals the tenant's reservation position, and the landlord's absolute and relative share of the surplus of the project is rising in the tenant's wealth.

Figure 7 illustrates this theorem. Allocative efficiency rises with the wealth of the tenant because on efficiency grounds the payment gradient $(h - l)$ should equal the expected marginal product of effort $(H - L)$, but where

[18] This is because placing two constraints upon the landlord's choice cannot improve the landlord's payoff over the single-constraint alternatives available to him outside the range $[\mathbf{w}_*, \mathbf{w}^*]$.

[19] To see that h rises, we differentiate $d'(e_0(\mathbf{w})) - \mathbf{w}$ with respect to $\mathbf{w}$, getting $d''e_0' - 1 = 1/e_0 - 1 > 0$.

the maximum penalty the landlord can impose on the tenant in the bad state is small, setting the payment gradient equal to the expected marginal product of effort can be accomplished only by raising h not by lowering l; because under these circumstances raising h not only enhances efficiency but distributes some of the efficiency gains to the tenant, the landlord's private incentives fail to implement the social optimum.

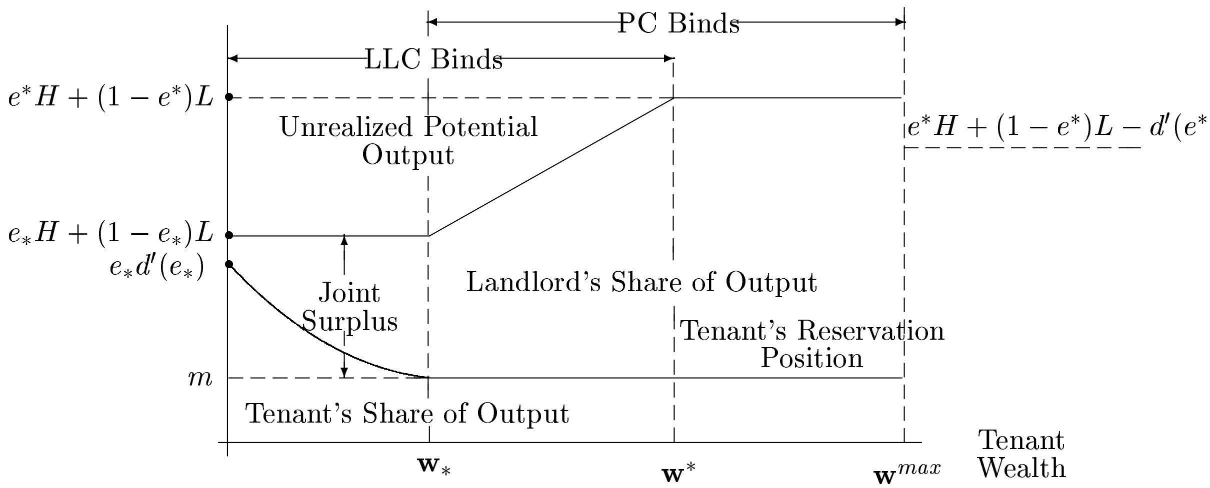

Figure 7: Tenant Wealth and Allocative Inefficiency. Efficient contracts require $\mathbf{w} \geq \mathbf{w}^*$. Lower wealth levels entail sub-optimal effort and expected output and (for $\mathbf{w} < \mathbf{w}_*$) a smaller tenant share of the joint surplus.

The intuition supporting Theorem (8), namely that tenant wealth attenuates allocative inefficiencies arising from contractual incompleteness, suggests that wealth limits will be more stringent where the noncontractible input effect is particularly powerful in inducing the good outcome, namely where $H - L$ is large, motivating

Theorem 9. Wealth Constraints and Noncontractible Inputs *The wealth limits $\mathbf{w}_*$ and $\mathbf{w}^*$ defined in Theorem 8 vary with the importance of effort in determining outputs, namely, with $H - L$.*

The reason, which may be confirmed from equations (11), (16), and (17), is that as $H - L$ increases, optimal effort is greater so the Pareto-optimal payment gradient is steeper and hence the LLC is binding for higher levels of tenant wealth.

Finally, we have

Theorem 10. Productivity Enhancing Redistributions to Tenants *For the contract defined in Theorem 8, a redistribution of wealth from the landlord to a tenant with wealth* $\mathbf{w}' \in [\mathbf{w}_*, \mathbf{w}^*)$ *would enhance allocative efficiency.*

The reasoning, which may be confirmed by inspection of equation (18), is that over this range of wealth tenant's expected utility is equal to $m + \mathbf{w}$ so an addition to tenant wealth benefits tenants directly, while the induced increase in their effort (namely $de/dw = 1/ed''(e)$) benefits the landlord, partially offsetting the cost of the transfer.

There is some wealth level $\mathbf{w}^{max}$ for which the land in question will be worth more to the tenant as owner than to the landlord, taking account of the available credit and other considerations discussed at the outset. As owner, the tenant will of course implement the surplus-maximizing effort level.

It is possible, indeed likely in many cases, that the tenant's fallback position m will not be independent of the tenant's wealth—as in the case, for example, where the tenant's wealth is held in the form of a truck. If m varies with $\mathbf{w}$, and if the wealth level $\mathbf{w}_*$ is defined as before but with $m = m(\mathbf{w})$ where $m'(\mathbf{w}) > 0$, the tenant's share of output will rise with tenant wealth for $\mathbf{w} > \mathbf{w}_*$, with the landlord's share unambiguously declining for $\mathbf{w} > \mathbf{w}^*$ and with the effect of increased tenant wealth on the landlord's share being ambiguous over the range $\mathbf{w}_*, \mathbf{w}^*$, the sign depending on the strength of the $m(\mathbf{w})$ relationship. For $\mathbf{w} < \mathbf{w}_*$, the landlord's share is increasing in $\mathbf{w}$, so the landlord would prefer wealthier tenants.

In the model here we have discussed only the one-period game. In a multi-period model there are interesting extra dimensions of the incentive effects under tenancy. The landlord may mitigate the problem of underapplication of the non-contractible input by using a threat of eviction when output is low.[20] The threat may be effective in our model with limited liability, since for $\mathbf{w} < \mathbf{w}_*$ the tenants earn some rent over and above their reservation income m, which they would lose if evicted. In a multi-period model, apart from labor effort (and other current input choice), the incentive to invest may also be affected by the eviction threat. Such a threat may discourage long-term improvements on the land (that are often non-contractible).[21] One may add that the removal of the threat (say, through

[20] This observation goes back to Johnson (1950) and has been formalized in Bardhan (1984) and Dutta, Ray and Sengupta (1989).

[21] Classical economists have emphasized this adverse effect of tenurial insecurity on investment. John Stuart Mill, for example, regarded this as the major defect of metayage

a land reform program that provides tenurial security) may also improve the bargaining power of the tenant, and investment may be encouraged because the tenant now expects to get a higher share of the additional output generated by that investment.

But as Banerjee and Ghatak (1996) point out, there are two ways in which eviction threats may also have a positive effect on the incentive to invest. First, just as eviction threats raise current labor effort because tenants care about the expected value of future rents from the work and the prospect of losing these rents induces them to work harder, similarly investment today raises the chances of doing well tomorrow and hence retaining the job day after tomorrow, thus they may respond positively to eviction threats. Secondly, if eviction threats raise current effort then that raises the chance of being around in the next period and this effect too is favorable to investment. All these put together make the net effect of eviction threats on investment rather ambiguous. The empirical evidence of Banerjee and Ghatak from a tenancy reform program in West Bengal (in the form of improvements in the security of tenure) suggests that the net effect of the program on the rate of growth of agricultural productivity is positive.

One major limitation of limited liability models in the context of rural areas in poor countries is that some of the main results are driven by the assumption that the asset-poor cannot be penalized enough by the landlord for rent default in bad times. Apart from the linearity of the utility function assumed, which is clearly implausible when consumption is near zero, in long-term relationships of closed village communities landlords can sometimes get around this problem with a weather-dependent side-payment (credit) to the tenant to be paid back in better times. For a model of such tenancy-cum-credit contracts, see Kotwal (1985). Historical information on agricultural production under different weather conditions as well as information on production on neighboring farms may be used by the landlord in the design of an incentive contract in such cases. It should also be kept in mind that in a traditional village context the landlord has access to various non-economic forms of punishing a defaulting poor tenant.

4 Wealth, Risk Aversion and Insurance

An important impediment to policy measures to redistribute economic resources in favor of nonwealthy producers is that to have the desired incentive effects, assigning residual claimancy to producers also involves assigning

in France. This effect is formalized in Bardhan (1984) and Banerjee and Ghatak (1996).

them control rights over the relevant assets. Nonwealthy producers, however, tend to be more risk averse than wealthy and/or highly diversified nonproducers—for instance stockholders, entrepreneurs, or landlords.[22] As a consequence, there is generally a tradeoff between effective production incentives and socially optimal risk choices. We explore ways of attenuating this tradeoff, extending an approach suggested by Domar and Musgrave (1944) and Sinn (1995).

A number of empirical investigations document a high level of risk aversion on the part of the nonwealthy. Low wealth entails lower return to independent agricultural production, for instance, because farmers sacrifice expected returns for more secure returns. Rosenzweig and Wolpin (1993) find that low-wealth Indian farmers seeking a means to secure more stable consumption streams, hold bullocks, which are a highly liquid form of capital, instead of buying pumps, which are illiquid but have high expected return. The relevant effects are not small. Rosenzweig and Binswanger (1993) find, for example, that a one standard deviation reduction in weather risk would raise average profits by about a third among farmers in the lowest wealth quartile (p. 75), and virtually not at all for the top wealth-holders. Moreover, they conclude that the demand for weather insurance would come primarily, if not exclusively, from poor farmers. Nerlove and Soedjiana (1996) find a similar effect in Indonesia with respect to sheep.[23]

Thus because of risk aversion, a reassignment of property rights to low-wealth producers might be unsustainable if as a result producers' income streams are subject to high levels of stochastic variation. Carter, Barham and Mesbah (1996) and Jarvis (1989) provide a vivid example: in the Central Valley of Chile three quarters of those families who received individual assignment of land rights under a land redistribution program in the 1970's sold their assets within a decade.

However, as Musgrave, Domar, and Sinn suggest, the availability of insurance can lead to increased risk-taking and willingness to hold risky assets. But the market for forms of insurance that promote risk-taking in production may be imperfect (Atkinson and Stiglitz 1980). Shiller (1993) provides several contemporary applications, arguing that capital market imperfections even in the most advanced economies lead to the absence of insurance markets for major sources of individual insecurity and inequality. For instance, a major form of wealth insecurity in many families is the capital value of the family home, due to medium- to long-term fluctuations in

[22] See Saha, Shumway and Talpaz (1994) and the many studies cite therein.

[23] See Hoff (1996a) for a discussion of this and related studies.

average housing prices in a region. No insurance for such fluctuations is available, but Shiller suggests that this and other similar insurance markets can be activated through proper financial interventions. Along these same lines, Sinn (1995) argues that the welfare state in the advanced economies can be understood in part as a successful set of policy measures to improve the risk-taking behavior of the nonwealthy where private 'social insurance' markets fail.

On the other hand, many attempts at preserving the small independent producer through extending credit availability and crop insurance have failed (Carter and Coles 1997), though these failures may be due to forms of insurance that are not incentive compatible (Newbery 1989). For instance, insuring individual crops reintroduces the same agency problems as sharecropping and wage labor. By contrast, as we show below, allowing producers to purchase insurance covering some general condition that is correlated with individual crop risk but that does not affect individual production incentives, can be effective in eliciting risk taking on the part of the nonwealthy without incurring efficiency losses. A crop insurance program in India, for example, based payments to individual farmers not on the output of their own plots but rather on average crop yields in larger agro-climatic regions to which they belong (Dandekar 1985). Disaster insurance for crops in the United States is similarly designed (Williams, Carriker, Barnaby and Harper 1993). Or insurance payments may be based on the exogenous source of the risk itself, if this is measurable. An example of this is rainfall insurance, whereby the producer pays a fixed premium, and receives a schedule of returns depending upon the average rainfall in the region over the growing season.[24]

In this spirit, we show below that under plausible conditions reducing the exposure of the nonwealthy to stochastic fluctuations independent of their productive activities can induce increased risk taking in production, and hence can help sustain otherwise unsustainable asset redistributions. General social insurance can also allow access to credit markets for wealth-poor agents who would be otherwise excluded. Platteau, Murickan, Palatty and Delbar (1980), Sanderatne (1986), Ardington and Lund (1995) and Deaton and Case (1997) provide some evidence for this phenomenon.

The model developed below shows that, exposed to the risk associated

[24]Similarly, the taxation of agricultural income can be based on general growing conditions rather than measured farm output, thus combining insurance and revenue-producing goals. The idea is not new. The *Zabt* system of taxation, developed by the Mughal rulers of North India during the Sixteenth Century, based assessments on estimates of the productive capacities of the land rather than on actual harvests (Richards 1993):85ff.

with residual claimancy, asset-poor producers

- may avoid buying projects that they could operate productively, even when they are financially capable of doing so, may sell rather than operate such projects that are transferred to them, and will choose suboptimal levels of risk for any project that they do retain and operate;
- there exists a class of productivity enhancing egalitarian asset redistributions that are sustainable as competitive equilibria but will not occur through private contracting even when loans are available to all producers at the risk-free interest rate;
- this class may be expanded by a offering fair insurance to nonwealthy asset holders that protects the producer against risk unassociated with the production process (e.g., health insurance, consumer goods price stabilization) or that protects independent producers against 'industry risk' that is unrelated to the quality of their own decisions;
- while competitive profit maximizing insurers may supply some forms of insurance of this type, they will generally do so in a suboptimal manner.

Let us begin with a risk neutral employer who owns an asset and employs a worker. The worker receives a wage w, and the project uses non-depreciable capital goods with value $\mathbf{k}$. We assume the employer must supervise the worker to guarantee performance, with supervision costs $m > 0$. We also assume the project consists of a continuum of possible technologies of varying risk and expected return, with higher risk yielding higher expected return over some range. We summarize the choice of technology in an expected net revenue schedule $g(\sigma)$, which is a concave function of the standard deviation of revenue $\sigma > 0$, with a maximum at some $\sigma^* > 0$.[25]

We then write the employer's profits, net of the opportunity costs of capital, $p(\sigma)$ as

$$p(\sigma) = \sigma z + g(\sigma) - \rho\mathbf{k} - m - w \tag{19}$$

[25] This shape follows from two plausible assumptions. First, production techniques that offer positive expected return involve a strictly positive level of risk. Hence expected return is an increasing function of risk for low levels of risk. Second, firms have access to production techniques that have very high returns when successful, but with a low probability of success (e.g. a firm may lower costs by not diversifying its product line, or by assuming the availability of particular production inputs). Hence above a certain point expected return declines with increasing risk.

where z is a random variable with mean zero and standard deviation unity and ρ is the risk-free interest rate.

The employer, who is risk neutral, maximizes $\mathbf{E}p(\sigma)$, the expected value of profits, giving first order condition

$$(\mathbf{E}p)_\sigma = g'(\sigma) = 0, \tag{20}$$

determining the expected profit-maximizing risk level σ^*. We further assume that the project is part of a competitive system with free entry, so profits must be zero in equilibrium. Since the employer is risk-neutral, this means the equilibrium wage rate w^* is given by

$$w^* = g(\sigma^*) - \rho\mathbf{k} - m. \tag{21}$$

Suppose the wage-earner considers becoming an independent producer by renting capital and undertaking production. To abstract from problems of credit availability, we assume that the productive equipment constituting the asset may be rented at a per-period cost $\rho\mathbf{k}$ where ρ is the risk-free interest rate. This is equivalent to assuming that the producer can borrow funds to purchase the asset at the risk-free rate. The independent producer's net payoff is then given by

$$y(\sigma) = \sigma z + g(\sigma) - \rho\mathbf{k}, \tag{22}$$

since being self-employed, the producer pays neither the wage nor the monitoring cost (we assume that the effort level of the producer remains the same). Indeed, the fact that the producer does not incur the monitoring cost captures our assumption that productive efficiency improves when the producer ceases being a wage-earner and becomes the residual claimant.

Suppose the producer has utility function $u(\mathbf{w})$, which is twice differentiable, increasing, and concave in wealth $\mathbf{w}$, and define

$$v(\sigma,\mu) = \mathbf{E}u(\mathbf{w}) = \int_{-\infty}^{\infty} u(\mu+\sigma z)dF(z), \tag{23}$$

where $F(z)$ is the cumulative distribution of z. Thus $v(\sigma,\mu)$ is the expected utility of the payoff $\mu+\sigma z$. We write the slope of the level curves $v(\sigma,\mu)=\bar{v}$ where $\bar{v}\in\mathbf{R}$.

$$s(\sigma,\mu) = -\frac{v_\sigma}{v_\mu}, \tag{24}$$

and we write the Arrow-Pratt risk coefficient for the agent as

$$\lambda(\mathbf{w}) = -\frac{u''(\mathbf{w})}{u'(\mathbf{w})}.$$

Then following Meyer (1987) and Sinn (1990), we know that $v(\sigma, \mu)$ behaves like a utility function where μ is a 'good' and σ is a 'bad.' The level curves $v(\sigma, \mu) = \bar{v}$ are indifference curves which, in the case of decreasing absolute risk aversion, are increasing, convex, flat at $\sigma = 0$, become flatter for increasing μ when $\sigma > 0$, and become steeper for increasing σ. Movements to the north and to the west thus indicate both improved welfare and flatter indifference curves. These properties are illustrated in Figure 8. We henceforth assume the producer exhibits decreasing absolute risk aversion, which means $\lambda'(\mathbf{w}) < 0$; i.e., the producer becomes less risk averse as wealth increases.[26]

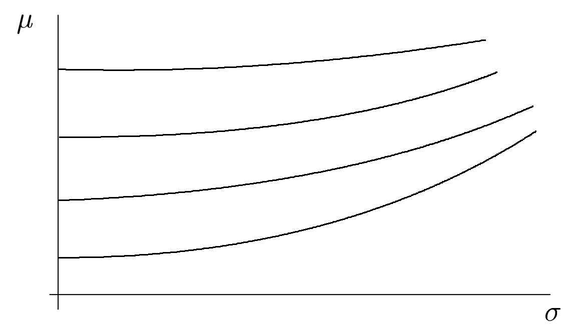

Figure 8: Indifference Curves of the Decreasingly Absolutely Risk Averse Producer with Utility Function $v(\sigma, \mu)$

The producer then chooses σ to maximize

$$\pi(\sigma) \equiv v(\sigma, \mu(\sigma))$$

where

$$\mu(\sigma) \equiv \mathbf{E}y(\sigma) = g(\sigma) - \rho\mathbf{k}, \tag{25}$$

giving the first order condition

$$\pi_\sigma = v_\mu[g'(\sigma) - s(\sigma, \mu(\sigma))] = 0. \tag{26}$$

This indicates that the marginal rate of transformation of risk into expected payoffs, $g'(\sigma)$, must equal the marginal rate of substitution between risk and expected payoff, $s(\sigma, \mu)$. The producer's optimizing problem as residual claimant is depicted in Figure 9 as choosing the highest indifference curve of $v(\sigma, \mu)$ that satisfies the constraint (25), which is just the tangency point at

[26] Virtually all empirical studies support decreasing absolute risk aversion. For a recent review of the literature, see Saha et al. (1994).

A, giving σ^o, which satisfies the first order condition (26). The producer's risk aversion implies $s(\sigma,\mu) > 0$, which by (26) requires that $\sigma^o < \sigma^*$, so the independent producer chooses a lower level of risk than the risk neutral employer.

The tradeoff between the allocative gains and suboptimal risk losses that occur when the asset is assigned to the asset-poor producer is illustrated in Figure 9. This figure depicts both the pre-transfer allocation in which the employer chooses σ^* and pays w^*, and the post-asset-transfer situation indicated by point A. The allocative gain associated with the transfer is the increase in the expected return from w^* to the point D, or just m, the saving in monitoring input. The suboptimal risk loss is $D - F$, reflecting the fact that the risk averse producer prefers point A to point C on the risk-return schedule. There is no reason, of course, to expect the gains to exceed the costs.

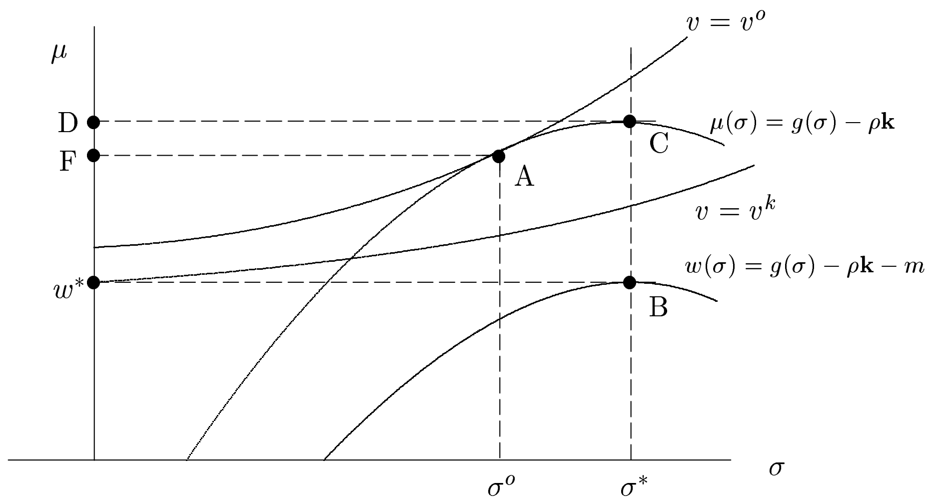

Figure 9: The Tradeoff between Gain in Expected Return and Lost in Suboptimal Risk Taking

To compare the producer's welfare as residual claimant as opposed to wage-earner, note that when the employer chooses σ, by (21) the equilibrium payoff to the employee occurs at the maximum point σ^* of the schedule $w = g(\sigma) - \rho\mathbf{k} - m$, as shown at point B in Figure 10(a). Figure 10(a) shows the case where the producer is better off as residual claimant rather than working for the employer, since the indifference curve through $(\sigma^o, \mu(\sigma^o))$ is higher than the indifference curve through $(0, w^*)$. By contrast, Figure 10(b) shows the case where the producer is better off working for the employer.

Notice that in this case the producer has higher expected income as residual claimant than as wage-earner, but is exposed to an excessive level of risk. The differences between the two cases is the greater degree of risk aversion assumed in the second case, as is indicated by the steeper indifference locus. Competitive equilibrium for the first case implies that the producer acquire the asset and in the second that the producer work for the employer, so in both cases the competitive assignment of residual claimancy and control rights would appear to implement an efficient solution.

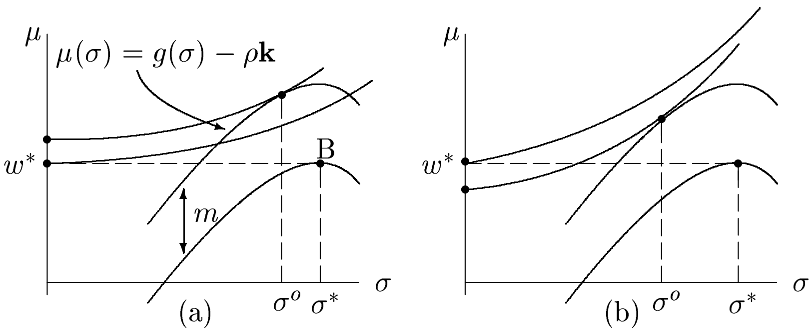

Figure 10: Comparing Wage Earning and Independent Production
Note that in (a) the producer is better off as residual claimant, and in (b) the reverse is true.

We now consider whether the analysis would be altered by an outright transfer of $\mathbf{k}$ to the employee, thus obviating the need to rent these assets. It might well be thought that the result would not change, as the producer's per period return from selling the asset $\rho\mathbf{k}$ is exactly the rental cost, so the asset transfer simply converts a direct cost (the cost of renting the capital) into an opportunity cost (the forgone cost of renting the capital to another agent), seemingly leaving the analysis unaffected. But this inference is unwarranted. Suppose the producer has wealth $\mathbf{w}$ independent from participation in production, and earns a secure income $\rho\mathbf{w}$ on this wealth. Then we find that

Theorem 11. *If the producer satisfies decreasing absolute risk aversion, the level of risk the producer assumes is an increasing function of wealth* $\mathbf{w}$.

This occurs because increasing the producer's wealth flattens the indifference curves in σ-μ space, so the optimal production point moves closer to the maximum on the risk-return schedule. To prove this formally, note that with wealth $\mathbf{w}$, (25) now becomes

$$\mu(\sigma) = \rho(\mathbf{w} - \mathbf{k}) + g(\sigma),$$

and the producer as before chooses σ to maximize $\pi(\sigma) \equiv v(\sigma, \mu(\sigma))$, giving the first order condition (26), which we totally differentiate with respect to $\mathbf{w}$ to obtain

$$\pi_{\sigma\sigma}\frac{d\sigma}{d\mathbf{w}} + \pi_{\sigma\mathbf{w}} = 0.$$

Now $\pi_{\sigma\sigma} < 0$ by the second order condition, and

$$\pi_{\sigma\mathbf{w}} = -\rho v_{\mu} s_{\mu} > 0,$$

since $s_{\mu}(\sigma_{\mathbf{w}}, \mu_{\mathbf{w}}) < 0$. Thus $d\sigma/d\mathbf{w} > 0$.

It follows that there exist wealth transfers of the following form: before the transfer, the producer prefers to work for an owner whose capital stock is $\mathbf{k}$. When an amount $\mathbf{k}$ of wealth is transferred to the producer, indifference curves become flatter, and in the new situation holding the productive asset and becoming an independent producer is the preferred alternative. The transfer is productivity enhancing because the increase in technical efficiency (elimination of m) is not offset by the output losses occasioned by the suboptimal risk level.

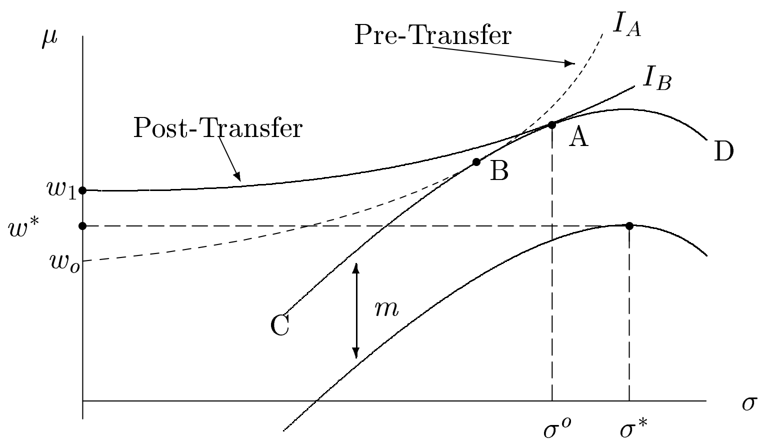

Figure 11: Example of a Productivity Enhancing Asset Redistribution

This is illustrated in Figure 11. In this figure, the before-transfer indifference curves for the agent are the dashed curves. Clearly wage labor dominates independent production. After the transfer, indicated by the solid curves, the decrease in risk aversion of the agent renders independent production superior to wage labor.[27] Thus the transfer is sustainable. This

[27]The utility levels corresponding to the dashed and the solid indifference curves are of course not the same. In particular, the dashed indifference curve through point $(0, w^*)$

result also demonstrates that the gains to the producer are sufficient to compensate the previous owner of the asset as the producer's returns to holding the asset exceed the opportunity cost $\rho\mathbf{k}$, which is identical to the required compensation.

An egalitarian wealth transfer may thus be productivity enhancing, although the compensation that rendered the transaction a Pareto improvement is not generally implementable, since a lump sum wealth transfer $\mathbf{k}$ to the former owner (or equivalently, an enforceable commitment of the producer to pay $\rho\mathbf{k}$ per period) would simply induce the producer to sell rather than operate the asset.

Credit market constraints played no part in this demonstration, as the producer was assumed to be able to borrow at the competitive risk-free interest rate ρ. However if the asset poor do face credit constraints insofar as a transfer of wealth may alleviate these constraints a second class of productivity enhancing asset transfers may exist. To see this assume that the cost of borrowing to the producer is $r(\mathbf{w})$ where $\mathbf{w} \geq 0$ is the total collateralizable wealth of the producer, where

$$r'(\mathbf{w}) < 0 \qquad \text{and} \lim_{\mathbf{w}\to\infty} r(\mathbf{w}) = \rho. \tag{27}$$

It is simple to show (as we do in Bowles and Gintis (1998c)) that if a credit constrained worker with wealth $\mathbf{w}$ faces an interest rate $r(\mathbf{w})$ satisfying (27), and a fraction κ of the value $\mathbf{k}$ of the capital requirements of the project can serve as collateral on a loan, then for sufficiently large $\mathbf{k}$ the transfer of the capital good to the agent is productivity enhancing.

To see this why this is so, note that a producer with wealth $\mathbf{w}$ can acquire the capital good at per period cost of $r(\mathbf{w}-(1-\kappa)\mathbf{k})\mathbf{k}$. The expected income μ_p for the work who purchases the asset is

$$\mu_p(\sigma) = g(\sigma) - r(\mathbf{w} - (1-\kappa)\mathbf{k})\mathbf{k},$$

while the expected income μ_t for the producer who has acquired the asset by transfer is

$$\mu_t(\sigma) = g(\sigma) - r(\mathbf{w} + \kappa\mathbf{k})\mathbf{k}.$$

Choose σ_p^* and σ_t^* to maximize $v(\sigma, \mu_p(\sigma))$, and $v(\sigma, \mu_t(\sigma))$, respectively. The producer who is employed and receiving the wage w^* would not benefit from purchasing the asset if $v(\sigma_p^*, \mu_p(\sigma_p^*)) < v^*(0, w^*)$, which is clearly true for sufficiently large $\mathbf{k}$. The same producer having received the asset $\mathbf{k}$ by

corresponds to a lower utility level than the solid indifference curve through $(0, w^*)$, since in the latter case the agent has higher wealth.

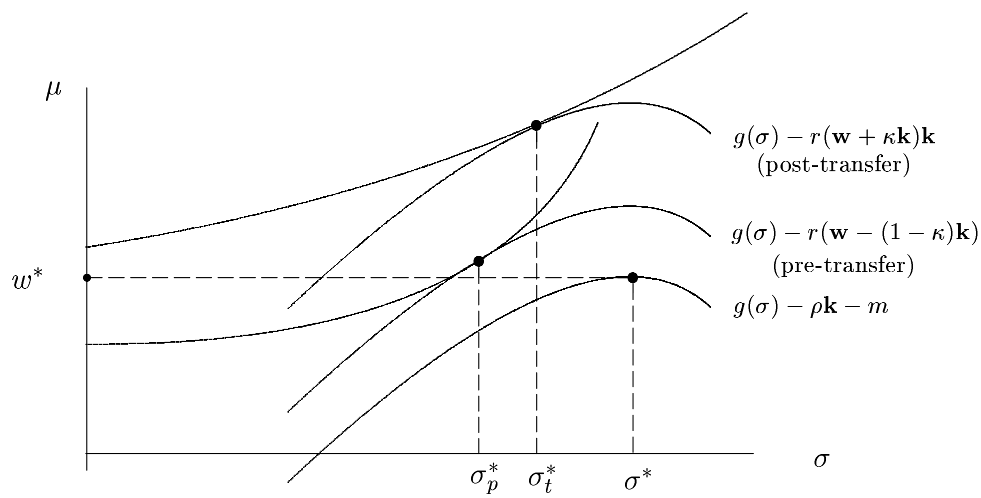

Figure 12: A Productivity Enhancing Redistribution where the Producer Faces a Credit Constraint

transfer would prefer to hold the asset if $v(\mu_t(\sigma_t^*), \sigma_t^*) > v*(0, w*)$. A productivity enhancing asset transfer thus requires that:

$$v(\sigma_p^*, \mu_p(\sigma_p^*)) < v*(0, w^*) < v(\mu_t(\sigma_t^*), \sigma_t^*).$$

Suppose the first inequality is satisfied. Since

$$\mu_t(\sigma) - \mu_p(\sigma) = [r(\mathbf{w} - (1-\kappa)\mathbf{k}) - r(\mathbf{w} + \kappa\mathbf{k})]\mathbf{k} > 0,$$

it is clear that, for sufficiently large $\mathbf{k}$, the second inequality will be satisfied as well at $\sigma = \sigma_p^*$, and hence *a fortiori* at $\sigma = \sigma_t^*$.

Thus where a wealth transfer will alleviate the credit market constraints faced by the wealth poor, productivity enhancing redistributions may exist even were the producers' risk aversion unaffected by the transfer. Hence wealth related credit constraints and wealth related risk aversion provide the basis for productivity enhancing asset redistributions. The two mechanisms are analogous in that in both cases the transfer of the asset reduces the costs associated with the assignment of residual claimancy and control rights to the wealth poor, attenuating suboptimal risk taking and the costs of risk exposure in the first, and reducing the opportunity cost of ownership in the second.

Figure 12 illustrates a productivity enhancing redistribution to the credit constrained wealth poor producer.

It follows that measures that render the producer less risk averse, or lessen the risk involved in production, lessen the risk allocation losses associated with the reassignment of residual claimancy and control rights to low-wealth producers. A producer who acquires the productive asset through an egalitarian redistribution policy, but who would otherwise prefer to sell this asset, could be induced by such measures to remain residual claimant on the use of the asset. In addition, such measures would reduce the losses from risk avoidance by producers willingly engaged in independent production. We shall suggest two plausible measures of this type. The first involves insuring producers against forms of risk exogenous to the production process, and the second involves insuring producers against public risk—risk correlated with the risk of independent production, but which is publicly observable.

Suppose the producer's wealth independent from participating in production, $\mathbf{w}$, has a stochastic element $\gamma\zeta$ of mean zero distributed independently from z, where $\gamma > 0$ is a constant. We call such a stochastic element *exogenous risk*, and we term a reduction in γ a *reduction in exogenous risk* (as opposed to the endogenous risk σz that chosen by the producer). In Bowles and Gintis (1998c) we show that under these conditions an economic policy measure that reduces the degree of uncertainty facing producers unrelated to the productive asset itself, for instance health insurance, consumer goods price stabilization, or business cycle stabilization, may induce non-wealthy producers to assume a higher level of risk exposure in production and thus increase the scope of application of productivity enhancing egalitarian redistributions.

A second measure with similar effects is insurance against public risk. Suppose the random variable η is positively correlated with the stochastic element z in production, and is publicly observable at the end of the production period, hence is contractible. We call η a *production-related public risk.* Average rainfall in the region over the growing season, for instance, is a form of production-related public risk. Consider a market for a fair insurance policy on production-related public risk that pays producers a premium l and obliges the producer to pay back an amount $b\eta$ at the end of the production period. We call this a *public risk insurance policy*, and we call b the *payback rate.* In Bowles and Gintis (1998c) we show that

Theorem 12. *There is a fair public insurance policy that induces the socially optimal level of risk-taking. Such a policy will not be purchased on a competitive insurance market, in which the producer can choose the payback rate to maximize expected payoff.*

The intuition underlying this result is that the socially optimal insurance policy induces risk neutral behavior by restricting the producer's choice to no insurance at all or more insurance than the producer would choose in a competitive environment. Profit maximizing producers would demand a lower level of insurance. The reason for the difference is that only when the degree of risk and the public signal are perfectly correlated does the insurance policy that renders the standard deviation of income invariant to the choice of risk level by the producer (inducing risk neutral behavior by the producer) also minimize the standard deviation of income (corresponding to the producer's desired fair insurance policy).

There are three reasons why the market in public risk insurance may fail. First the competitively determined insurance rate does not achieve the socially optimal outcome. Second, the market in public risk insurance is subject to adverse selection if $r_{z\eta}$ differs among producers and is not public knowledge. Third, a private industry selling public risk insurance may not be able to operate as approximately risk neutral, since the signal η is a macroeconomic variable that is perfectly correlated for all insurance purchasers, so insurance companies cannot use the law of large numbers to handle the volatility of their payouts. Moreover, if there is uncertainty concerning μ_η, or if μ_η shifts over time, the insurance companies' risk position becomes even more precarious. Thus government policy might be needed to implement this outcome.

Of course an analysis of the defects of the market solution to the independent producer's risk problem must be complemented by an analysis of the defects of the public sector as an insurance provider. In particular, in the absence of a mechanism guaranteeing their accountability, public decision-makers will choose the level and type of independent producer insurance to meet multiple objectives, of which fostering socially efficient production is only one.

5 Wealth Constraints and Residual Claimancy in Team Production

Monitoring by peers in work teams, credit associations, partnerships, local commons situations, and residential neighborhoods is often an effective means of attenuating incentive problems that arise where individual actions affecting the well being of others are not subject to enforceable contracts (Whyte 1955, Homans 1961, Ostrom 1990, Tilly 1981, Hossain 1988, Dong and Dow 1993b, Sampson, Raudenbush and Earls 1997). Most explanations of the incentives to engage in mutual monitoring (Varian 1990, Stiglitz 1993) rely either on the small size of the interacting group, or on repeated interactions and low discount rates, allowing the Folk theorem to be invoked. Neither of these is completely satisfactory, since work teams are often large and the Folk theorem has little explanatory power.[28] Other treatments leave the incentive to engage in mutual monitoring unexplained (Arnott 1991, Weissing and Ostrom 1991).[29]

We provide an explanation of mutual monitoring in single shot interactions among members of large teams. The key conditions supporting mutual monitoring are (a) the fact that when members are residual claimants, shirking imposes costs on other team members, and contributing to production becomes a team norm, and (b) a fraction of team members are 'reciprocators' who punish violators of team norms. We will provide evidence that under appropriate conditions reciprocators occur with sufficient frequency to sustain cooperative outcomes.

The problem of free riding in teams has been addressed by two standard models. The first, due to Alchian and Demsetz (1972), holds that residual claimancy should be assigned to an individual designated to monitor team members' inputs, thus ensuring the incentive compatibility for the (non-contractible) activity of monitoring itself, while addressing the members' incentive to free ride by the threat of dismissal by the monitor. They contrast this view of the 'classical firm,' as they call it, with an alternative in which team members are residual claimants and monitoring is performed, if at all, by salaried personnel. Alchian and Demsetz correctly observe that group residual claimancy would dilute incentives, but simply posit the allocational superiority of the classical firm: "we assume that if profit sharing had to be relied on for all team members, losses from the resulting increase in central monitor shirking would exceed the output gains from the increased incentives of other team members not to shirk." (1972):786 As we will see,

[28] The repeated game solution to the problem of sustaining cooperative behavior in teams has several weaknesses, including: (a) there are a multiplicity of equilibria, most of which do not exhibit high levels of cooperation; (b) subgame perfection (i.e., the credibility of threats to punish non-cooperators) requires an implausible degree of coordination among team members.

[29] Dong and Dow (1993b) and Legros and Matthews (1993) assume the team can impose collective sanctions on shirkers. This assumption is reasonable if shirking is easily detected and team members have more effective or lower cost forms of punishment than are available to a traditional firm. We do not make this assumption. Dong and Dow (1993a) assume shirking can be controlled by the threat of non-shirkers to exit the team. However the threat of exiting is credible only if team members have very high fallback positions—in Dong and Dow's model, this takes the form of independent production—which generally is not the case.

their invocation of the so-called "$1/n$ problem" to justify this assumption is not entirely adequate. Moreover, the classical firm they describe is not an accurate description of the way that conventional firms—except for the very smallest—handle the problem of monitoring, for monitoring is commonly done by salaried supervisors rather than residual claimants.

The second approach, pioneered by Hölmstrom (1982), demonstrates that in principal multi-agent models one can achieve efficiency or near-efficiency through contracts that make individual team members residual claimants on the effects of their actions without conferring ownership rights on them. Contracts of this type typically impose large penalties for shirking and require large lump-sum up-front payments on the part of agents, or they pay each team member the entire team output minus a large constant and thus, in the presence of stochastic influences on output, entail negative payments in some periods, or at best a substantial variance of income to team members. These arrangments are infeasible if team members have insufficient wealth. Moreover, where contributions (e.g., work effort) are continuously variable these incentive mechanisms support large numbers of Nash equilibria, thus rendering breakdown of cooperation likely.

These approaches do not explain how mutual monitoring works, but rather why it may be unnecessary. The limited applicability of the owner-monitor and optimal contracting approaches provides one motivation for exploring the relationship between residual claimancy and mutual monitoring in teams. Another motivation is empirical. There is some evidence that group residual claimancy is effective, by comparison with payments unrelated to group output, even in quite large teams (Hansen 1997). Mutual monitoring based on residual claimancy appears to be effective in the regulation of common pool resources such as fisheries, irrigation, and grazing lands (Ostrom 1990), in the regulation of work effort in producer cooperatives (Greenberg 1986, Craig and Pencavel 1995) and in the enforcement of non-collateralized credit contracts (Hossain 1988). Experimental studies (Frohlich, Godard, Oppenheimer and Starke 1997) provide additional support for the effects of residual claimancy in inducing lower supervision costs and higher productivity in (small) work teams. Further, the fact that residual claimancy may provide incentives for monitoring even in quite complex settings and large groups is suggested by evidence that in the United States home ownership is a significant predictor of participation in community organizations (Glaeser and DiPasquale forthcoming) and local politics but, significantly, not national politics (Verba, Schlozman and Brady 1995), as well as willingness to monitor and sanction coresidents who transgress social norms (Sampson et al. 1997).

Locating residual claimancy in teams can have positive incentive effects, since team members may have privileged access to information concerning the activities of other team members, and may have means of disciplining shirkers and rewarding hard work that are not available to third parties. As residual claimants, moreover, team members may have the incentive to use this information and exercise their sanctioning power, even if the team is large. Thus while Alchian and Demsetz are surely correct in saying that residual claimancy in large teams does not substantially reduce the direct incentive to free ride, it may support superior means of sanctioning and hence discouraging free riding through mutual monitoring.[30] Monitoring is costly, however, and if the desire to monitor is not sufficiently widespread, we shall see, mutual monitoring will fail.

We will show that under certain conditions, residual claimancy by team members can provide sufficient incentives for mutual monitoring, and thus support high levels of team performance. A key element in our approach, one shared by recent contributions of Kandel and Lazear (1992), Rotemburg (1994), Banerjee et al. (1994), and Besley and Coate (1995) is that our model is based on 'social preferences' which, while unconventional, are well supported by recent experimental and other research.

We assume that though team members observe one another in their productive activity, they cannot design enforceable contracts on actions because this information is not verifiable (cannot be used in courts). In this situation we show that under appropriate conditions the assignment of residual claimancy to team members will attenuate incentive problems even when team size is large.

Two common characteristics of successful mutual monitoring are uncontroversial: the superior information concerning non-verifiable actions of team members available to other team members and the role of residual claimancy in motivating members to acquire and use this information in ways that enhance productivity. Less clear is whether residual claimancy motivates costly monitoring in large groups.[31]

[30] Some models of mutual monitoring are presented in Varian (1990), Kandel and Lazear (1992), Weissing and Ostrom (1991), Dong and Dow (1993a,b), and Banerjee, Besley and Guinnane (1994). Other models of incentives in teams include Hölmstrom (1982), McAfee and McMillan (1991), Legros and Matthews (1993), Rotemburg (1994), and Besley and Coate (1995)

[31] The problem of motivating the peer-monitors would not arise, of course, if team members were sufficiently altruistic towards teammates. In this case members would simply internalize the benefits conferred on others by their monitoring. Rotemburg (1994) develops a model of this type. More generally, Robert Frank writes: "Under [profit sharing] plans, the injury caused by an act of shirking affects not only the shareholders of the firm

A parsimonious explanation of mutual monitoring is provided, however, by the notion of *strong reciprocity*: the well-documented human propensity to cooperate with those who obey, and to punish those who violate social norms, even when this behavior cannot be justified in terms of self-regarding, outcome-oriented preferences (Campbell 1983, Bowles and Gintis 1998a). We distinguish this from *weak reciprocity*, namely reciprocal altruism, tit-for-tat, exchange under complete contracting, and other forms of mutually beneficial cooperation that can be accounted for in terms of self-regarding outcome-oriented preferences. The commonly observed rejection of substantial positive offers in experimental ultimatum games is consistent with this interpretation.[32] Moreover the fact that offers generated by a computer rather than another person are significantly less likely to be rejected suggests that those rejected offers at to cost to themselves are reacting to violations of norms rather than simply rejecting disadvantageous offers (Blount 1995). More directly analogous to the team production case, however, are findings in n-player public goods experiments. These provide a motivational foundation for mutual monitoring in teams whose members are residual claimants, since these experiments show that agents are willing to incur a cost to punish those whom they perceive to have treated them or a group to which they belong badly.[33] In these experiments, which allow subjects to punish non-cooperators at a cost to themselves, the moderate levels of contribution typically observed in early play tend to rise in subsequent rounds to near the maximal level, rather than declining to insubstantial levels as in the case where no punishment is permitted. It is also significant that in the experiments of Fehr and Gächter, punishment levels are undiminished in the final rounds, suggesting that disciplining norm violators is an end in itself and hence will be exhibited even when there is no prospect of modifying the subsequent behavior of the shirker or potential

but also the shirker's co-workers. Individual workers who care about their co-workers will be reluctant to impose these costs... even when it is impossible for co-workers to observe the act of shirking." (1991):168. However were team members sufficiently altruistic in this sense to motivate mutual monitoring, there would be no initial free rider problem either.

[32]See Güth, Schmittberger and Schwarz (1982), Ostrom, Walker and Gardner (1992), Güth and Ockenfels (1993), Forsythe, Horowitz, Savin and Sefton (1994), Cameron (1995), Hoffman, McCabe and Smith (April, 1996), and Falk and Fischbacher (1998). For an overview of the studies in this area, see Davis and Holt (1993) and Fehr, Gächter and Kirchsteiger (1997).

[33]See Ostrom et al. (1992) on common pool resources, Fehr et al. (1997) on efficiency wages, and Fehr and Gächter (1996) on public goods. Coleman (1988) develops the parallel point that free riding in social networks can be avoided if network members provide positive rewards for cooperating.

future shirkers.

The willingness to engage in costly punishment provides a basis for linking residual claimancy with mutual monitoring, even in large teams. An individual who shirks inflicts harm on the other members of the team if (and only if) they are residual claimants. Members may then see this violation of reciprocity as reason to punish the shirker. We should note that our model requires only that a certain fraction of team members be reciprocators. This is in line with the evidence from experimental economics, which indicates that in virtually every experimental setting a certain fraction of the subjects do not retaliate, either because they are self-interested, or they are purely altruistic.[34]

To see how mututal monitoring works, consider a team with n members ($n > 3$), each of whom can either Work, supplying one unit of effort, or Shirk, supplying zero units of effort. We assume the members of the team are equal residual claimants on team output, but there may be other residual claimants outside the team (e.g., equity-holders who do not engage in production, or a government that taxes output). For convenience, if we refer to i or j, we assume they are team members in $\{1,\ldots,n\}$ unless otherwise stated, and if we refer to both i and j, we assume $i \neq j$. Also we write $n_{-i} = \{k = 1,\ldots,n | k \neq i\}$. We assume agents have linear utility functions that are additive in costs and benefits.

Let σ_j be the probability that member j shirks, so $\sigma = \sum_{j=1}^{n} \sigma_j/n$ is the average rate of shirking. The value of team output net of nonlabor costs is the number of workers working times q, the average (and marginal) net product of effort, which we can write as $n(1-\sigma)q$. Each member's payoff is then given by $(1-\sigma)\alpha q$, where $\alpha \in [0,1]$ is defined as the team's residual share. The loss to the team from one member shirking is αq. The gain to an individual from shirking is the disutility of effort, $b > 0$, which we assume is identical for all team members. We also assume $\alpha q > b$, otherwise universal shirking would be optimal.[35]

Consider a single team member j. Another member $i \in n_{-j}$ can either Monitor j at cost $c_i > 0$, or Trust j at zero cost. Member i imposes a cost $s_i > 0$ on j if i detects j shirking. This cost may involve public criticism, shunning, threats of physical harm and the like. We assume that acts of punishment, like work effort, are non-verifiable and hence not subject to contract. If j shirks and i monitors j, we assume the shirking will be detected

[34] For an especially clear example, see Blount (1995). Fehr and Schmidt (1997) provides a survey of rejection rates in ultimatum games.

[35] Most of the homogeneity assumptions we make can be dropped, at the expense of complicating the notation and the descriptions of the model.

with probability $p_{ij} \in (0,1]$, where this *probability of detection* may vary with the spatial proximity of team members and the transparency of the production process, and other factors that we do not model here.

We model the incentive to monitor by supposing that i experiences a subjective gain $\rho_i(\alpha) \geq 0$ from disciplining a shirking member j, which occurs if j shirks, i monitors j, and j is detected.[36] The harm done by a shirker to team members is proportional to the degree of residual claimancy so it is reasonable to assume that the strength of the norm of cooperation increases with the degree of residual claimancy of the team. Thus $\rho_i(0) = 0$ and $\rho_i'(\alpha) \geq 0$ for all $i = 1, \ldots, n$. We call $\rho_i(\alpha)$ i's *propensity to punish shirkers*.[37] Note that some members may exhibit no propensity to punish; i.e., $\rho_i(\alpha) = 0$, $\alpha \in [0,1]$.

If i monitors, the likelihood of detecting j shirking is $\sigma_j p_{ij}$, so the net cost of monitoring j over trusting j is $c_i - \sigma_j p_{ij} \rho_i(\alpha)$. Then if μ_{ij}, $i \in n_{-j}$, the probability that i monitors j, is chosen to be a best response, we have

$$\mu_{ij} \begin{cases} = 0, & c_i > \sigma_j p_{ij} \rho_i(\alpha) \\ \in (0,1), & c_i = \sigma_j p_{ij} \rho_i(\alpha) \\ = 1, & c_i < \sigma_j p_{ij} \rho_i(\alpha) \end{cases} \tag{28}$$

Let s_j^* be the expected punishment inflicted by all $i \in n_{-j}$ on j if j shirks. We have

$$s_j^* = \sum_{i \in n_{-j}} p_{ij} \mu_{ij} s_i. \tag{29}$$

Writing the direct gain to an individual from shirking as

$$g = b - \frac{\alpha q}{n}, \tag{30}$$

the expected gain to j from shirking, including the expected cost of punish-

[36] In fact, ee expect ρ_i to depend on αq, but since we do not vary q in our analysis, we suppress q in the argument. Note that, unlike a member's share of the firm's net revenue, the subjective gain from punishing does not decline with the size of the team. We motivate this assumption and discuss the effects of team size below. We assume $\rho_i(\alpha) < s_j$ for all $\alpha \in [0,1]$, to avoid bizarre '*sado-masochistic*' optima in which team members cooperate by shirking, punishing, and being punished, the net psychic return to which is greater than the return to working.

[37] For simplicity we have assumed that a monitor's propensity to punish, ρ_i, is not affected by the propensities to punish or the observed rates of punishing of other members of the team. Replacing this with the assumption that punishing propensities are positively related opens the possibility of multiple equilibria, some involving high levels of punishing and some low. We explore this alternative below.

ment, is $g - s_j^*$. Therefore if σ_j is chosen as a best response, we have

$$\sigma_j \begin{cases} = 0, & g < s_j^* \\ \in [0,1], & g = s_j^* \\ = 1, & g > s_j^* \end{cases} \tag{31}$$

If $g < 0$ there is a unique, Pareto efficient, Nash equilibrium in which no members shirk and no member monitors.[38] In this case residual claimancy alone is sufficient to ensure efficiency. The more interesting case, however, is where group size is sufficiently large that residual claimancy alone does not entail incentive compatibility. We thus suppose in the rest of the paper that $b > q/n$, so that even with full residual claimancy assigned to the team as a whole, shirking is an individual best response in the absence of monitoring. In this case any Nash equilibrium involves positive shirking, since if $\sigma_j = 0$ for some j then by (28) $\mu_{ij} = 0$ for $i \in n_{-j}$. But then by (31), $\sigma_j = 1$, a contradiction. Thus we must investigate conditions under which $0 < \sigma_j < 1$ for some j in equilibrium, requiring

$$g = s_j^*. \tag{32}$$

We call such a situation a *working equilibrium.*

We say i is a *reciprocator* if $\rho_i(\alpha) > 0$. Suppose the fraction of reciprocators is f, so the remaining fraction $1 - f$ of team members are self-regarding—for these agents $\rho_i(\alpha) = 0$, and they never monitor or punish. Notice that if j is not a reciprocator, j has fn potential monitors, whereas if j is a reciprocator, j has $fn - 1$ potential monitors. In the a interest of simplicity of exposition, we will ignore this difference, assuming all agents face fn potential monitors.[39]

We say reciprocators are *homogeneous* if there are *parameters* p, c, s and $\rho(\alpha)$ such that if i is a reciprocator, then $c_i = c$, $p_{ij} = p$, $s_i = s$, and $\rho_i(\alpha) = \rho(\alpha)$. We have

Theorem 13. *Suppose reciprocators are homogeneous, with parameters* p, c, s *and* $\rho(\alpha)$. *If* $c > p\rho(\alpha)$, *then the unique Nash equilibrium satisfies* $\sigma_j = \sigma^* = 1$ *and* $\mu_{ij} = \mu^* = 0$ *for all* $i, j = 1, \ldots, n$; *i.e., all members shirk and no member monitors. If* $c < p\rho(\alpha)$ *and* $g > fnps$ *the unique Nash equilibrium satisfies* $\sigma_j = \sigma^* = \mu^* = 1$ *for all* $i, j = 1, \ldots, n$; *i.e., all workers shirk and all members monitor. If* $c < p\rho(\alpha)$ *and* $g < fnps$, *then*

[38] This equilibrium is efficient because we have ruled out 'sado-masochistic' optima.

[39] The effect of dropping this assumption is in all cases quite transparent, since in effect, the model is the union of n independent games, in each of which one agent is the worker and the other $n - 1$ agents are the monitors.

(a) *there is a mixed strategy Nash equilibrium in which all members shirk with probability*

$$\sigma^* = \frac{c}{p\rho(\alpha)} \tag{33}$$

and reciprocators monitor with probability

$$\mu^* = \frac{g}{fnps}; \tag{34}$$

(b) *the effect of residual claimancy on the incidence of shirking,* $\partial\sigma^*/\partial\alpha$, *is independent of team size and the fraction of reciprocators.*

(c) *the social welfare difference per team member between a first best world with no shirking and the equilibrium of this game is* σq. *This is declining in the degree of residual claimancy and is independent of team size and the fraction of reciprocators.*

We note that while the fraction of reciprocators and the level of punishment they may inflict do not appear in (33), they are not unimportant in the determination of the level of shirking, for if f and s are sufficiently small, the condition $g < fnps$ is violated, and universal shirking occurs.[40] Notice also that, as one would expect, shirking declines with an increase in the propensity to punish shirkers, an increase in the probability of detecting shirking, or a decrease in the cost of monitoring.

The intuition behind part (a) of Theorem 13 is as follows. The equilibrium level of shirking, σ^*, equates the net benefits of monitoring and trusting, while the equilibrium level of monitoring equates the net benefits of working and shirking. Thus when σ is greater than its equilibrium value (33), the expected benefits of monitoring, $\sigma p\rho$, exceed the costs c. Members who monitor with high probability will then receive higher payoffs than members who monitor with low probability, inducing some to increase their monitoring probability. As the monitoring probability increases, the gains to shirking decline, leading suppliers to reduce σ. This dynamic continues until (33) is satisfied. A similar dynamic occurs when σ is less than its equilibrium value. When μ is greater than its equilibrium value (34), the expected costs of shirking $f\mu nps$ exceed the benefits g. Suppliers who work with high probability will then be receiving higher payoffs than suppliers who shirk with high probability, inducing some to decrease their rate of shirking. As

[40] We have not investigated equilibria when $c = p\rho(\alpha)$ or $g = fnps$. This is because these cases are neither generic nor interesting.

the shirking rate declines, the gains to monitoring decline, leading to a reduction in μ. This dynamic will continue until (34) is satisfied. A similar dynamic occurs when μ is less than the equilibrium value given by (34).

Behavioral traits such as a work ethic or a willingness to punish co-members for inflicting harm on the team are, of course, strongly norm governed and as such need not be proximately determined by the explicit optimization of any agent but rather may be the expression of behavioral rules. Thus the model underlying Theorem 13 may be interpreted as the basis of a dynamic treatment of work and punishment norms, with the updating of norms responding to the observed payoffs of others. For example, as our description of the intuition behind part (a) of Theorem 13 suggests, the determination of σ and μ may be represented as dynamic processes based on the differential replication of norms governing the working, shirking, monitoring and punishing behaviors we have modeled, the equilibrium values σ^* and μ^* simply representing outcomes that are stationary in the underlying dynamic. We do not develop this extension here.

It may be objected that it is plausible to treat $\rho(\alpha)$ as a decreasing function of team size, on the grounds that strong reciprocity may weaken when the team becomes larger, and thus the propensity to punish any given act of shirking would fall. Though this is possible, there is to our knowledge no clear evidence in support of this notion, and there are many 'stylized facts' contradicting it. For instance, people are often observed to support their local sports team, their regional sports team, and their national sports team with equal commitment.

There are of course additional paths through which increasing team size might weaken the mutual monitoring mechanism. Increased n might lower the cost s a monitor can impose on a shirker, since the 'average social distance' between a pair of workers can be expected to increase as the team becomes more numerous. The ability to detect shirking may also decline. It is clear from (33) that lowering p will indeed reduce the efficiency of the team, while (34) shows both mechanisms lead to an increase in the monitoring level required to prevent shirking. Since a change in μ does not affect the efficiency of the system, we will investigate only the former effect.

To model the relationship between team size and detection probability in a plausible manner, we will drop our homogeneity assumption. In its place we will assume that for any two team members i and j, either i can or cannot see j. Suppose that if j shirks and i inspects j, the probability of i detecting j shirking is p if i j, and is zero otherwise. Suppose in all other respects, the model is as described above. Then if there is an integer $k \geq 1$ such that as team size $n \geq k$ increases, each member sees exactly k

other team members, then if $fkps > g$, our previous assertions hold with k substituted for n in the denominator of (34). The intuition behind this result is that as long as increasing team size does not reduce the number of team members that one may 'see' the effectiveness of mutual monitoring does not decline as team size increases.

However large teams often do not have the informational homogeneity assumed above, since with increased size often comes a more refined division of labor in which there are specialized 'work groups' whose members all see one another, and who are not seen by other team members. Members of such groups have an incentive to collude by agreeing that the reciprocators in the group will not monitor, and hence all are free to shirk without penalty. We call this a 'shirking clique.' Of course in a one-shot situation the promise not to monitor and punish is not credible, but it can be supported in a repeated game framework as a form of 'tacit collusion.' Teams can reduce the frequency of such behavior by rotating members among work groups, rendering the effective discount rate for the repeated game too high to support collusion, but such rotation may entail prohibitive organizational costs.

We conclude that under appropriate conditions, strong reciprocity can operate even in large teams, so that the allocative efficiency case for residual claimancy, and hence asset holdings by team members, is not necessarily weakened, unless the frequency of reciprocators is too low or the division of labor favors the widespread formation of shirking cliques.

If residual claimancy provides motives for mutual monitoring, the distribution of wealth may have allocational effects as in the more commonly treated cases of concerning human investment, agrarian tenancy, and entrepreneurship (Loury 1981, Galor and Zeira 1993, Laffont and Matoussi 1995, Bowles and Gintis 1998b). The reason is that some distributions (those in which team members are without wealth, for example) effectively preclude the assignment of residual claimancy to team members, because transferring residual claimancy over the income streams of an asset but not ownership itself to team members creates incentives for the team to depreciate the assets, the costs of which may more than offset any gains from mutual monitoring. Thus prohibitive costs may arise if residual claimancy is separated from ownership, and outright ownership may be precluded by borrowing limitations faced by zero wealth team members.

Thus a redistribution of wealth to team members may improve the allocational efficiency of the team. If, as in the case modelled, team members are risk neutral, such a redistribution must be potentially Pareto-improving, in the sense that the gains of beneficiaries of the redistribution exceed the amount necessary to fully compensate the losses of those whose wealth was redistributed. Such potentially Pareto improving asset redistributions to team members are necessarily egalitarian, assets being provided to the wealth poor. The reason is that in the absence of borrowing limitations imposed by low wealth, the allocational gains associated with residual claimancy *via* mutual monitoring would be fully exploited by voluntary transactions among agents. Thus wealthy team members would not be precluded from becoming residual claimants and mutual monitors were this to improve allocational efficiency.

Where team members are not only poor but risk averse as well (the more probable case) this felicitous efficiency equity complementarity (rather than trade off) is still possible, but less likely.[41] The reason is that control rights over the use of the assets must accompany residual claimancy and ownership and, as we have seen in the previous section, risk averse team members would likely implement a socially suboptimal level of risk taking in decisions concerning investment and technical choice.

In Section 2 we surveyed the advantages of locating residual claimancy and hence asset ownership with agents making noncontractible production-related decisions with the resulting allocative efficiency prescription that where all agents are risk neutral those who control noncontractible actions should also own the results of those actions, thus requiring producers to hold the relevant assets. In this section we have asked under what conditions the logic of this prescription might be extended from the case of individual producers to team producers. We found that the mechanism of reciprocal fairness can operate even in large teams, so that even in large teams the allocative efficiency case for residual claimancy, and hence asset holdings by team members, is not necessarily weakened.

We saw in Section 2 that nonwealthy producers may be precluded from acquiring ownership of productive assets, and this obviously precludes nonwealthy members of productive teams from acquiring the ownership needed to render mutual monitoring effective. One may then ask if the competitively determined assignment of residual claimancy and control are inefficient in the sense that a productivity enhancing redistribution of assets of the type defined in the introduction might be possible. This question cannot be adequately addressed without dropping the assumption of general risk neutrality. Fortunately the results of Section 4 may be readily extended to the case of team production, suggesting that a reallocation of ownership to team members may be productivity enhancing even when the members

[41] We explore this case in Bowles and Gintis (1998c).

are risk averse.

6 Initial Asset Inequality and Cooperation on the Local Commons

One of the important, though somewhat neglected, ways in which asset inequality can affect economic efficiency is through influencing the likelihood of cooperation in the management of local public goods, varying from personal security and neighborhood amenities in urban residential settings to local commons situations. In particular, the daily livelihood of vast masses of the rural poor in many countries depends on the success with which common-pool resources (CPRs)—such as forests, grazing lands, in-shore fisheries, and irrigation water—are managed, and the environmental consequences of their management. A CPR is defined by Ostrom (1990), p. 30, as "a natural or man-made resource system that is sufficiently large as to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from its use." There are several documented examples of successful local-level cooperation on CPRs in different parts of the world—see Ostrom (1990) for many such examples—but there are also numerous cases of failure of cooperation. Understanding the factors that lead to success or failure of community management of these resources is critical.

CPR management is a collective action dilemma: a situation in which mutual cooperation is collectively rational for a group as a whole, but individual cooperation may not be rational for each member. One factor that has not always been recognized as critical to the outcome of collective action dilemmas is heterogeneity among the players. Here our attention will, of course, be largely restricted to a single but potent kind of heterogeneity: asset inequality.

Olson (1965) hypothesized that inequality may be beneficial to the provision of a public good:

> In smaller groups marked by considerable degrees of inequality—that is, in groups of members of unequal 'size' or extent of interest in the collective good—there is the greatest likelihood that a collective good will be provided; for the greater the interest in the collective good of any single member, the greater the likelihood that member will get such a significant proportion of the total benefit from the collective good that he will gain from seeing that the good is provided, even if he has to pay all of the cost himself. (p. 34)

Inequality in this context can thus facilitate the provision of the collective good, with the small players free-riding on the contribution of the large player. Supporting this position, Bergstrom, Blume and Varian (1986) show that, in a very general setting, wealth redistributions that increase the wealth of a positive contributor to a public good will lead the latter to demand more (and therefore entail an increase the supply) of the public good.[42] These analyses of the supply of public goods are relevant to conservation among CPR users. Restraint in resource use is analytically equivalent to contributing to a public good. Following these studies, we would expect group heterogeneity to be conducive to the effective management of CPRs.

Nevertheless, field studies of CPR management have often shown that the relationship between inequality and collective action is more complex. Bardhan (1995) reviews the case study literature regarding the relationship between inequality and cooperation in locally-managed irrigation systems, primarily in Asia, and finds that while equality tends to favor successful local management in many cases, highly unequal traditional authority structures may also have served the efficient husbanding of resources. Baland and Platteau (1997) likewise summarize many relevant examples from the case-study literature; they focus more on forests, fisheries and grazing lands, and on African cases. In an econometric study of the determinants of cooperation in 104 local peasant committees in Paraguay, Molinas (1998) shows that there is a u-shaped relationship between inequality of land distribution and cooperative performance (in activities that include CPR management and provision of local public goods). With recourse to both theoretical analysis and empirical examples, Baland and Platteau show that inequality in resource-use entitlements has an ambiguous impact on the efficiency of the equilibrium outcome in the completely unregulated case. In regulated settings, inequality does not as a rule make common property regulation easier.

In the rest of this section we provide some simple and general theoretical arguments to analyze the effect of asset inequality on cooperation within a group, drawing upon the model in Dayton-Johnson and Bardhan (1997). Although the model is couched in terms of a fishery, the qualitative results should be in principle transferable to the case of other CPRs. In a two-player noncooperative model we explore the necessary and sufficient conditions for resource conservation to be a Nash equilibrium, and we show that,

[42] Chan, Mestelman, Moir and Muller (1996) report that when the Bergstrom-Blume-Varian model is tested in the laboratory, it correctly predicts the direction (though not the magnitude) of change in group contributions when income is redistributed toward positive contributors. It does not do so well in predicting individual behavior: individuals with low incomes overcontribute to the public good, and high income individuals undercontribute.

- contrary to the implication of the Olson hypothesis, increasing inequality does not, in general, favor full conservation. However,
- once inequality is sufficiently great, further inequality may push the players closer to efficiency.

Thus the theoretical model can generate a u-shaped relationship between inequality and economic efficiency.

In the preceding sections we have focused on a structure of incentives and constraints arising from a deliberately chosen contract. In contrast in this section the governing structure could be thought of more as a norm. The norm is viewed as self-sustaining if it is a Nash equilibrium in a simultaneous move game. The game is one of complete information. Markets are, however, incomplete because it is difficult to restrict access to the commons (whereas market incompleteness in previous sections arises because some actions that affect the gains from exchange are private information).

Consider a lake in which fish are born in the Spring and are mature in the Fall. If two fishers have access to the lake, it is then efficient for them to cooperate in waiting until the Fall to harvest the fish. However it may pay one of the fishers to defect and harvest in the Spring, when there is no competition from the other fisher. Knowing this, it may be profitable for the second fisher to do the same. Moreover, if a nonwealthy fisher cannot compete effectively against a wealthy fisher, it may never pay the nonwealthy fisher to cooperate. Thus the efficient output may require a considerable degree of wealth equality. On the other hand with great wealth inequality, the efficient output can be approximated, simply because the poor fisher has the means to harvest only a small portion of the stock of fish in the Spring.

Suppose the fishers are $i = 1, 2$, and each fisher i is endowed with wealth $\mathbf{w}^i > 0$, representing the fishing capacity of the fisher's capital goods (boats, tackle, nets, and the like), measured in number of units of fish that can be obtained in a period. There are two periods $t = 1, 2$ (Spring and Fall), in each of which fisher i can apply some or all of his fishing capacity to farming the lake.

Let F be the stock of fish in the lake. In the first period, fisher i must choose to use some portion $\mathbf{k}^i$ of his capacity $\mathbf{w}^i$ in fishing, so $\mathbf{k}^i \le \mathbf{w}^i$. Fishing yield is then given by

$$\phi^i = \begin{cases} \mathbf{k}^i & \text{for } \mathbf{k}^1 + \mathbf{k}^2 \le F; \\ \frac{\mathbf{k}^i F}{\mathbf{k}^1 + \mathbf{k}^2} & \text{otherwise.} \end{cases}$$

i.e., if the total take is less than F, fishers can fish to their capacity. Otherwise the fishers share F in proportion to the respective capacities they have applied to farming the lake.

Between Spring and Fall the stock of fish grows at rate $g > 0$, so that in period 2 the supply of fish is $(1+g)(F - \phi^1 - \phi^2)$.[43] We assume the future is not discounted, so each fisher's utility is simply the total amount of fish he catches. Clearly then in any efficient outcome there will be no fishing in period 1.

In the second period, each fisher i again chooses to apply some portion of his capacity $\mathbf{w}^i$ to farming the lake. We make the following *commons dilemma assumption*:

$$\mathbf{w} \equiv \mathbf{w}^1 + \mathbf{w}^2 \geq (1+g)F. \tag{35}$$

This assumption insures that the threat of resource degradation is sufficiently acute. Alternatively, (35) can be interpreted as a feasibility condition: the fishers are capable of harvesting the entire stock if they desire.

In the subgame consisting of the second period, both fishers will always fish to capacity, and will receive second period payoff

$$(1+g)(F - \phi^1 - \phi^2)\frac{\mathbf{w}^i}{\mathbf{w}},$$

where $\mathbf{w} = \mathbf{w}^1 + \mathbf{w}^2$ is the total wealth of the two fishers. However one fisher's period 1 action will enter the other fisher's period 2 payoff, and vice-versa. Thus we must concentrate on the fishers' actions in the first period. A strategy for each fisher is just a capacity choice $\mathbf{k}^i$ in the first period, so a strategy for the game is a pair $\{\mathbf{k}^1, \mathbf{k}^2\}$. We have

Theorem 14. *The strategy profile {0,0} in which neither fisher harvests in the first period is a Pareto optimum.*

We call this situation a *first best.*

The goal of conservation in fisheries is to reduce fishing to some level so that the remaining stock at the end of every period is sufficient to guarantee the survival of the fish population. In our simple model, that level has been normalized to zero in the first period. The second period extends to the end of the fishers' relevant economic horizons.[44]

[43] If g were negative there would be no real dilemma. First period depletion of the resource would be an equilibrium and an optimum.

[44] In this model we have abstracted from the problem of discount rates in order to focus

The following theorem notes the conditions under which the least efficient outcome is a Nash equilibrium.[45]

Theorem 15. *If* $\mathbf{w}^i > \frac{g}{1+g}F$ *for* $i = 1, 2$, *then* $\{\mathbf{w}^1, \mathbf{w}^2\}$ *(i.e., complete resource depletion) is a Nash equilibrium.*

If the inequality in Theorem 15 holds, each fisher has sufficient capacity that if one fishes in the Spring, so little will remain in the Fall that the other fisher's best response is not to wait but also to fish in the Spring, sharing the undepleted catch. Note in particular that Theorem 15 holds when $\mathbf{w}^i > F$ for $i = 1, 2$, so each fisher can unilaterally harvest the whole lake. When is full conservation a Nash equilibrium?

Theorem 16. *Full conservation is an equilibrium of the model if and only if*

$$\mathbf{w}^i \geq \frac{\mathbf{w}}{1+g} \qquad \textit{for } i = 1, 2.$$

To see why this is true, suppose fisher 2 conserves. Then for every unit fisher 1 harvests in the first period, he gives up $(1+g)\mathbf{w}^1/\mathbf{w}$ in the second period. Thus $(1+g)\mathbf{w}^1/\mathbf{w} \geq 1$ is the threshold above which fisher 1 will conserve, conditional on conservation on the part of fisher 2.

This theorem suggests the following corollary. Let

$$\Delta(\mathbf{w}) = \{(\mathbf{w}^{1'}, \mathbf{w}^{2'}) | \mathbf{w}^{1'}, \mathbf{w}^{2'} \geq 0 \text{ and } \mathbf{w}^{1'} + \mathbf{w}^{2'} = \mathbf{w}\}$$

be the set of all distributions of $\mathbf{w}$. We say $(\mathbf{w}^{1'}, \mathbf{w}^{2'}) \in \Delta(\mathbf{w})$ is a *mean preserving spread* of $\{\mathbf{w}^1, \mathbf{w}^2\}$ if $|\mathbf{w}^{1'} - \mathbf{w}^{2'}| > |\mathbf{w}^1 - \mathbf{w}^2|$. We have

Corollary 16.1. *If* $(\mathbf{w}^{1'}, \mathbf{w}^{2'})$ *is a mean preserving spread of* $(\mathbf{w}^1, \mathbf{w}^2)$, *then full conservation is an equilibrium with initial wealth* $(\mathbf{w}^{1'}, \mathbf{w}^{2'})$ *only if it is an equilibrium with initial wealth* $(\mathbf{w}^1, \mathbf{w}^2)$. *Also, for all* $(\mathbf{w}^1, \mathbf{w}^2) \in \Delta(\mathbf{w})$ *there is a mean preserving spread* $(\mathbf{w}^{1'}, \mathbf{w}^{2'})$ *such that full conservation is not an equilibrium with wealth* $(\mathbf{w}^{1'}, \mathbf{w}^{2'})$.

more clearly on the incentives to conserve a resource. Formally, the discount rate would be subtracted from g, the rate of resource regeneration. If the discount rate is greater than g, first period depletion of the fishery is optimal, and conservation is not economically rational. Furthermore, as we have seen above, it is reasonable to suppose that each fisher's discount rate is a decreasing function of wealth. In this case, the more unequal the distribution endowments, the more difficult it will be to sustain universal conservation of the resource. Taking account of a poor fisher's high rate of time preference is equivalent to the situation in which the poor fisher faces a low rate of growth of the stock and hence has little incentive to conserve.

[45]For proof of this and the subsequent theorems in this section, see Dayton-Johnson and Bardhan (1997).

The Olson hypothesis that inequality enhances the prospects for collective action can be interpreted as a comparative static statement: increasing inequality for a given level of aggregate wealth makes full conservation more likely. The Corollary above suggests that this is not so. The second part of the Corollary states that, starting from any wealth distribution, there exists a less equal wealth distribution such that full conservation is not an equilibrium. In particular, if full conservation is an equilibrium under the initial distribution, then we know from Theorem 16 that $\mathbf{w}^i \geq \mathbf{w}/(1+g)$ for $i = 1, 2$. Then wealth can be taken from one fisher until $\mathbf{w}^i < \mathbf{w}/(1+g)$ for that fisher. Hence full conservation is no longer an equilibrium.

The Corollary to Theorem 16 shows that increased inequality does not necessarily lead to equilibrium conservation. Theorem 17, shows that under maximum inequality—that is, when one fisher holds all of the wealth—conservation is an equilibrium.

Theorem 17. *If $g \geq 0$ then under perfect inequality ($\mathbf{w}^1 = 0$ or $\mathbf{w}^2 = 0$), full conservation is an equilibrium.*

In part, Theorem 17 reflects Olson's hypothesis that cooperation is more difficult in a group the larger the number of group members. In our fishery, conservation is an equilibrium outcome when the number of fishers with positive wealth is reduced to one. The above theorems consider only the conditions under which full conservation by both fishers is an equilibrium. The more realistic case in an unregulated fishery, and the case which may be closer to Olson's thinking, is the one in which changes in the distribution of wealth change the level of efficiency among a set of inefficient equilibria. This is considered in the following theorem. Theorem 18 says that if the distribution of wealth is sufficiently unequal already, then making it even more unequal can increase efficiency.

Theorem 18. *Define $F^*(\mathbf{w}^1, \mathbf{w}^2)$ as the minimum amount of first period fishing among all Nash equilibria of the game when the distribution of endowments is $(\mathbf{w}^1, \mathbf{w}^2)$. Whenever $\mathbf{w} > (1+g)F$, there exists $(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2) \in \Delta(\mathbf{w})$, such that for all mean preserving spreads $(\mathbf{w}^{1'}, \mathbf{w}^{2'})$ of $(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2)$, we have $F^*(\mathbf{w}^{1'}, \mathbf{w}^{2'}) < F^*(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2)$.*

Indeed, the proof of Theorem 18 demonstrates that for the wealth distribution $(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2) = (\mathbf{w} - gF/(1+g), gF/(1+g))$ and all mean preserving spreads of $(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2)$, fisher 1 will conserve regardless of the other's behavior. The theorem also illustrates that the full conservation equilibrium under perfect inequality in Theorem 17 is a limiting case as inequality is increased. For distributions such as $(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2)$, one fisher captures a sufficiently large share of the returns to conservation that he will unilaterally conserve. In particular, there exists an equilibrium in which the larger fisher conserves, the smaller fisher does not, and any mean preserving spread increases efficiency. If it were true that i's endowment were greater than $\mathbf{w}/(1+g)$, then by Theorem 16, fisher i will always conserve if fisher j does. If it were true that is endowment were greater than $\mathbf{w}/(1+g)$, then by Theorem 16, i would always conserve if j did. Thus any mean preserving spread of $(\hat{\mathbf{w}}^1, \hat{\mathbf{w}}^2)$, by reducing fisher i's capacity, will increase efficiency, since fisher j will play zero and more fishing will be deferred until the second period. This, then, is the commons analogue of the Olson public-goods hypothesis.

This situation is summarized in Figure 13, which shows (assuming $g > 1$, which is clearly necessary for a cooperative equilibrium in the two-person case examined here) that as fisher 2's share increases from 1/2, full efficiency is maintained until his share reaches $g/(1+g)$, at which point fisher 1 defects, reducing the total catch. Then as the share of fisher 2 continues to increase, the efficiency of the system increases apace, since fisher 1 is capable of harvesting a decreasing fraction of the fish stock in period 1. When fisher 2 owns all the wealth, full efficiency is restored.[46]

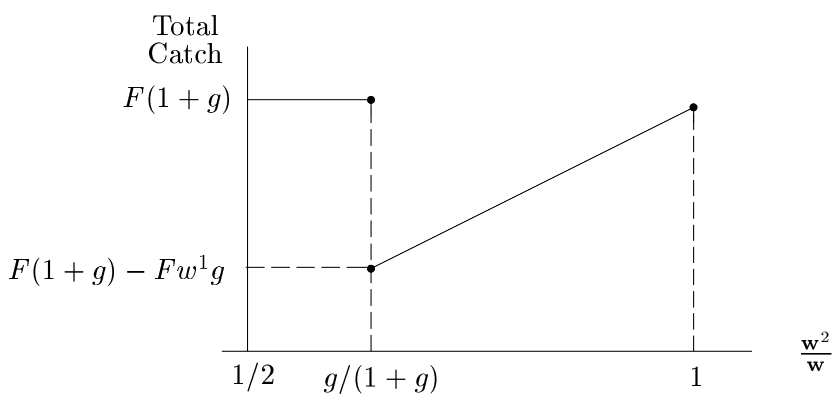

Figure 13: Inequality and the Efficiency of Cooperation

In fisheries worldwide, it has often been observed that large fishing companies with more opportunities to move their fleets elsewhere (compared to the small-scale local fishers) are much less concerned about conservation of fish resources in a given harvesting ground. This phenomenon of differential exit options depending on differential wealth levels extends also to

[46]This figure is due to J. M. Baland, personal communication.

other CPRs. On the other hand, there are cases where the poorer or smaller users may exercise an exit option.[47] In order to analyze such cases Dayton-Johnson and Bardhan (1997) extend the basic game presented above to the case when there is an exit option depending on a fisher's endowment level, $\mathbf{w}^i$; exit refers to investing or deploying one's capacity in another sector. In general, any comparative-static assertions about whether full conservation is a Nash equilibrium under different wealth distributions depends on the nature of the exit option function. In this connection Roland Bénabou has pointed out to us that, consistent with his discussion of 'inequality of income versus inequality of power' in his paper (1996), what matters is not inequality of wealth *per se*, but inequality of wealth relative to exit options and threats. If the value of one fisher's exit option grows faster than one-for-one with his wealth, then wealth inequality will foster rather than hinder cooperation.

If exit options are concave functions of wealth, increased wealth inequality does not, in general, enhance the prospects for full conservation. If full conservation is an equilibrium in a situation of perfect equality, then there is a mean preserving spread of the wealth distribution under which full conservation is not an equilibrium. Under the unequal distribution of wealth, it is the poorer fisher who finds it in his interest to play the exit strategy. But if the exit option functions are convex, it is the poorer fisher who has an interest in conditional conservation, while the richer fisher prefers the exit strategy. The nature of the exit option functions is ultimately an empirical question. In many situations, an exit option function could plausibly be linear beyond some level of wealth, but at lower levels of wealth it may be convex as a result of borrowing constraints.

The noncooperative model sketched above points to the nature of the complicated relationship between inequality and cooperation in an unregulated commons situation. One of the themes emphasized by many writers in the current policy discussion of the commons is that such problems are best described not always as prisoner's dilemmas, but rather that in many cases they may be problems of coordinating among multiple equilibria (Runge 1981, Ostrom 1990). This assertion is shared by our model: when the conditions of Theorem 16 are satisfied (i.e., $\mathbf{w}^i \geq \mathbf{w}/(1+g)$ for $i = 1, 2$), both resource degradation (depletion of the fish stock in period 1) and full conservation (no fishing in period 1) are equilibria. However, under many parameter configurations, the problem is indeed a prisoner's dilemma: full conservation, though a Pareto optimum, is not an equilibrium.

[47] For many examples of exit by the large and the small, see Baland and Platteau (1997).

One might presume that in real-world commons problems, economic actors often craft rules to regulate community use of common-pool resources. In the context of our non-cooperative model of a fishery we may discuss one possible regulatory mechanism that takes the form of asset redistribution: fishers may decide to redistribute wealth before the game is played in order to secure Pareto optimal outcomes.[48] One is then interested in knowing whether a first-best outcome can be realized, particularly in cases where a first-best is not an equilibrium outcome of the unregulated game.

Theorems 15 through 17 above and their corollaries, regarding the basic model, are comparative-static results considering the effect on efficiency of changes in the wealth distribution. If we make the assumption that wealth can be redistributed, these results can be reinterpreted as statements about the effects of redistribution. Thus, Theorem 16 tells us that, for asset distributions which give each fisher positive wealth, full conservation is an equilibrium if and only if each fisher's share of total wealth is greater than $1/(1+g)$. If g is at least one, then there always exists a wealth transfer (perhaps negative) from fisher 1 to fisher 2 such that full conservation is an equilibrium outcome. With the appropriate wealth transfer, full conservation can be supported as an equilibrium, even if it was impossible under the initial distribution. However, one may ask whether both fishers (in particular, the fisher who is asked to give up some wealth) would agree to such a transfer—or is this scheme of social regulation in fact Pareto optimal?

Let us say that the fisher who must cede some wealth to the other is fisher 2. If the fishers do not agree to transfer s between them, presumably the bad equilibrium will be played. In that case, fisher 2's payoff is $\mathbf{w}^2 F/\mathbf{w}$. If the transfer is effected and the good equilibrium results, fisher 2's payoff is $(\mathbf{w}^2 - s)(1+g)F/\mathbf{w}$. Is the latter greater than the former? It is, as long as

$$s < \frac{gw^2}{1+g}.$$

This condition on the size of the transfer is always satisfied if the condition in Theorem 16 is satisfied post-transfer for fisher i.

In this section we have focused on particular mechanisms linking wealth inequality and economic performance.[49] The case study literature refers to

[48] In line with our earlier sections on land and on team production, we are assuming that credit market imperfections inhibit the operation of a market in boats and other assets in achieving first-best results.

[49] Most of the economics literature concentrates on problems of *sharing costs* in collective action dilemmas. Elster (1989) argues that problems of *sharing benefits* may frequently lead to the breakdown of collective action. The latter problem is usually one of income

a much richer variety of such mechanisms. In particular, social norms can be powerful enforcers of cooperative agreements, but this power may be attenuated in extremely unequal environments. Individuals may observe some cooperative norms, but only in relation to the set of individuals they regard as their peers. This perspective has its roots in the theory of social exchange, one of whose founders, George Homans, comments: "The more cohesive a group... the greater the change that members can produce in the behavior of other members in the direction of rendering these activities more valuable." Public goods experiments by Kramer and Brewer (1984) eliciting levels of resource use in a commons tragedy situation found that common group identity (as opposed to within group heterogeneity) contributed strongly to conservation of the common resource.

This position receives some support from the experimental evidence. Indeed, contrary to many conventional treatments, bargaining agents often fail to reach the Pareto efficient bargaining frontier for reasons initially surveyed by (Johansen 1979). Initial inequality may be a cause of these bargaining failures. Socially beneficial cooperation often fails to materialize where the relevant actors cannot agree on and precommit to a division of the gains from cooperation. The resulting bargaining breakdowns are likely to occur where the bargaining power or wealth of the actors is particularly disparate. Experimental evidence suggests that subjects whose fallback positions are very different are less likely to come to agreements than are more equally situated subjects (Lawler and Yoon 1996). Further, the extent of cooperation and hence the average payoff in one-shot prisoner's dilemma games is inversely related to an experimentally manipulated social distance between the subjects (Kollock 1997). The resulting bargaining failures may occur because inequality heightens informational asymmetries among the bargaining partners, because very unequal offers based on disparities in initial wealth or bargaining power are likely to be perceived as unfair and rejected, as in experimental play of the ultimatum game (Camerer and Thaler 1995, Rabin 1993), or because changes in the rules of the game necessary to allow precommitments to *ex post* divisions of the gains to cooperation may be vetoed by the wealthy, who may fear the general redistributive potential of such institutional innovations.

It may also be that inequality affects the extent of enforceability of socially regulated solutions. The transaction costs for regulatory mechanisms may differ with the level of pre-existing inequality. These ideas are yet to

inequality rather than wealth inequality, although highly unequal initial asset distributions are likely to exacerbate the problem.

be formalized.

7 Conclusion

The study of incomplete contracts prompted two reconsiderations of the relationship between inequality and allocative efficiency. The first, a theoretical concern, has led many economists to reject the canon that allocational and distributional issues are separable and to recognize the stringency of the assumptions by which the Fundamental Theorem of Welfare Economics and the Coase theorem had initially established this separability result. The second, a more practical concern, has been to reconsider the policy relevance of the so called efficiency equality tradeoff. In the preceding pages we have surveyed some of the reasoning motivating both reconsiderations.

Where asymmetry or nonverifiability of information, or nonexcludability of users, makes contracts incomplete or unenforceable, and where for these and other reasons there are impediments to efficient bargaining, we have shown that private contracting will not generally assign the control of assets and the residual claimancy over income streams of projects to achieve socially efficient outcomes.

Can a mandated redistribution of wealth from the rich to those with few assets—perhaps in conjunction with other policies addressed to market failures arising from contractual incompleteness, for example in insurance—do better? We have seen that there are cases where such a redistribution will be sustainable in competitive equilibrium and will allow the nonwealthy to engage in productive projects that would otherwise not be undertaken, or to operate such projects in a more nearly socially optimal manner, or will support a more socially efficient use of common property resources. The subjective costs to the nonwealthy of increased risk exposure associated with residual claimancy on a variable income stream may be attenuated both by the effect of the asset transfer itself and by insurance against risks that are exogenous and public.

Thus mandated asset redistributions may rectify or attenuate the market failures resulting from contractual incompleteness. But is there any reason to expect that the indicated redistributions would be from rich to poor rather than the other way around? In the pages above we have mentioned three cases in which highly concentrated assets may contribute to allocative efficiency: in attenuating common pool resource problems, in providing incentives for the monitoring of managers, and in inducing a socially optimal level of risk taking. But in cases such as these, the asset will be worth more to the wealthy than to the nonwealthy, and private contracting alone is sufficient to ensure an efficient assignment or property rights, for the wealthy do not face the credit market disabilities that sometimes prevent the nonwealthy from acquiring residual claimancy and control rights. Thus the class of productivity enhancing asset redistributions is predominantly from the wealth to the asset poor.

The costs of mandated asset redistributions must also be carefully considered. The social welfare gains from a producitivity enhancing asset redistribution accrue to the recipients. As we have stressed, there may be no feasible means of recovering the costs of the redistribution from the recipients without destroying the incentive effects upon which the gains depend. Thus in general the government will be obliged to finance such redistribution by increasing its revenue through taxation and other fiscal means. The disincentive effects of such measures are well known and potentially severe (Buchanan, Tollison and Tullock 1980) but need not outweigh the allocative benefits of the redistribution (Hoff and Lyon 1995). Of course these disincentive costs fall on all forms of egalitarian redistribution, including health and unemployment insurance, income transfers and job creation programs, not just redistributions that are productivity enhancing.

It is clear, then, that distributional and allocational issues are thus inextricable, and that while there can be no presumption that egalitarian redistribution will improve efficiency, the conventional presumption to the contrary must also be rejected.

Recognition of the importance of incomplete contracts has had another consequence in economics: the revival of the classical economists' concern with "getting the institutions right." It is now commonplace to attribute national differences in economic performance to differences in institutional structures, and to explain persistent economic backwardness by institutions that fail to align incentives in productivity enhancing ways. Many economists equate "getting the institutions right" with establishing unambiguous property rights, along with institutions for the unimpeded transfer of these rights. In this they follow Coase (1960): "...all that matters (questions of equity aside) is that the rights of the various parties should be well defined..." (p. 19).

But Coase himself stressed the crucial nature of his assumption of zero transaction costs, so we can set aside as utopian the possibility that property rights could be perfected to such an extent that all external effects are internalized through complete contracting. If incomplete contracts are thus unavoidable, we have shown that "what matters" for allocative efficiency includes who holds the property rights and not simply that the rights be well defined.

The reasons extend considerably beyond the cases we have considered above, but they may be summarized as follows. We know that differing initial distributions of assets may persist over long periods. Further the pattern of holdings may exercise a powerful influence on the viability of differing structures of economic governance, by which we mean the entire nexus of formal and informal rules governing economic activities. The effects of wealth differences on patterns of residual claimancy and control that we have stressed are examples. But so is the more comprehensive sharp contrast in the institutional structure of societies with yeoman as opposed to latifundia based agriculture Engerman and Sokoloff (1994). While highly inefficient governance structures will not be favored in competition with substantially more efficient ones, the selection process is both slow and imperfect. Douglass North comments that "economic history is overwhelmingly a story of economies that failed to produce a set of economic rules of the game (with enforcement) that induce sustained economic growth." (1990):113.

Thus institutions may endure for long periods because they are favored by powerful groups for whom they secure distributional advantage. For this reason inequality in assets may impede economic performance by obstructing the evolution of productivity enhancing institutions. In addition to the incentive problems on which we have focused, this may be true both because maintaining highly unequal distributions of assets may be costly in terms of resources devoted to enforcing the rules of the game and because at least under modern conditions inequality may militate against the diffusion of cultural norms such as trust that are valued precisely because they are often able to attenuate the problems arising from contractual incompleteness.

References

Aghion, Philippe and Patrick Bolton, "A Theory of Trickle-down Growth and Development," *Review of Economic Studies* 64,2 (April 1997):151–172.

Alchian, Armen and Harold Demsetz, "Production, Information Costs, and Economic Organization," *American Economic Review* 62 (December 1972):777–795.

Alesina, Alberto and Roberto Perotti, "Income Redistribution, Political Instability, and Investment," *European Economic Review* 40 (1996):1203–1228.

Ardington, Elisabeth and Frances Lund, "Pensions and Development: Social Security as Complementary to Programs of Reconstruction and Development," *Southern Africa* 12,4 (August 1995):557–577.

Arnott, Richard, "Moral Hazard and Nonmarket Institutions," *American Economic Review* 81,1 (March 1991):180–190.

Atkinson, Anthony and Joseph Stiglitz, *Lectures on Public Economics* (New York: McGraw-Hill, 1980).

Baland, J. M. and J. P. Platteau, "Wealth Inequality and Efficiency in the Commons," 1997. CRED, University of Namur.

Banerjee, Abhijit and Andrew Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy* 101,2 (April 1993):274–298.

— and Maitreesh Ghatak, "Empowerment and Efficiency: The Economics of Tenancy Reform," 1996. Massachusetts Institute of Technology and Harvard University.

Banerjee, Abhijit V., Timothy Besley, and Timothy W. Guinnane, "Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test," *Quarterly Journal of Economics* (May 1994):491–515.

Bardhan, Pranab, *Land, Labor and Rural Poverty: Essays in Development Economics* (New York: Columbia University Press, 1984).

— , "Rational Fools and Cooperation in a Poor Hydraulic Economy," in Kaushik Basu et al. (ed.) *Choice, Welfare, and Development: A Festchrift in Honor of Amartya K. Sen* (Oxford: Clarendon Press, 1995).

Barham, Bradford, Steve Boucher, and Michael Carter, "Credit Constraints, Credit Unions, and Small-Scale Producers in Guatemala," *World Development* 24,5 (1996):792–805.

Barham, Vicky, Robin Boadway, Maurice Marchand, and Pierre Pestieau, "Education and the Poverty Trap," *European Economic Review* 39,7 (August 1995):1257–1275.

Barro, Robert, "Democracy and Growth," *Journal of Economic Growth* 1,1 (1996):1–28.

Bénabou, Roland, "Inequality and Growth," March 1996. NBER Macroeconomis Annual.

— , "What Levels of Redistribution Maximize Income and Efficiency?," March 1997. NYU mimeo.

Benhabib, Jess and Marc Speigel, "Cross Country Growth Regressions," May 1997. New York University.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian, "On the Private Provision of Public Goods," *Journal of Public Economics* 29 (1986):25–49.

Berry, Albert R. and William R. Cline, *Agrarian Structure and Productivity in Developing Countries* (Baltimore, MD: Johns Hopkins University Press, 1979).

Besley, Timothy and Stephen Coate, "Group Lending, Repayment Incentives and Social Collateral," *Journal of Development Economics* 46 (1995):1–18.

Binswanger, H. P., K. Deininger, and G. Feder, "Power, Distortions, Revolt and Reform in Agricultural and Land Relations," in *Handbook of Development Economics* Vol. 3B (Amsterdam: North-Holland, 1995).

Birdsall, Nancy, Thomas Pinckney, and Richard Sabot, "Inequality, Savings and Growth," 1996. Inter-American Development Bank, Office of the Chief Economist, Working Paper.

Black, Jane, David de Meza, and David Jeffreys, "House Prices, the Supply of Collateral and the Enterprise Economy," *Economic Journal* 106 (January 1996):60–75.

Blanchflower, David and Andrew Oswald, "What Makes a Young Entrepreneur?," *Journal of Labor Economics* (forthcoming).

Blount, Sally, "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences," *Organizational Behavior & Human Decision Processes* 63,2 (August 1995):131–144.

Bowles, Samuel and Herbert Gintis, "Credit Market Imperfections and the Incidence of Worker Owned Firms," *Metroeconomica* 45,3 (October 1994):209–223.

— and — , "The Evolution of Strong Reciprocity," 1998. Santa Fe Institute Working Paper.

— and — , "Mutual Monitoring in Teams: The Effects of Residual Claimancy and Reciprocity," 1998. Santa Fe Institute Working Paper.

— and — , "Risk Aversion, Insurance, and the Efficiency-Equality Tradeoff," 1998. University of Massachuetts Working Paper.

Buchanan, James, Robert Tollison, and Gordon Tullock, *Toward a Theory of the Rent-seeking Society* (College Station: Texas A&M University Press, 1980).

Camerer, Colin and Richard Thaler, "Anomalies: More Ultimatum and Dictator Games," *Journal of Economic Perspectives* 9,2 (1995):209–219.

Cameron, Lisa, "Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia," 1995. Discussion Paper #345, Department of Economics, Princeton University.

Campbell, Donald T., "Two Distinct Routes beyond Kin Selection to Ultra-Sociality: Implications for the Humanities and Social Sciences," in Diane L. Bridgeman (ed.) *The Nature of Prosocial Development* (New York: Academic Press, 1983) pp. 11–41.

Carter, Michael, Bradford Barham, and Dina Mesbah, "Agro Export Booms and the Rural Poor in Chile, Guatamala and Paraguay," *Latin American Research Review* 31,1 (1996):33–66.

Carter, Michael R. and Bradford Barham, "Level Playing Fields and *Laissez Faire*: Post-Liberal Development Strategies in Inegalitarian Agrarian Economies," *World Development* 24,7 (1996):1133–1150.

— and Dina Mesbah, "Can Land Market Reform Mitigate the Exclusionary Aspects of Rapid Agro-Export Growth?," *World Development* (July 1993).

— and Jonathan Coles, "Inequality-Reducing Growth in Latin American Agriculture: Towards a Market Friendly and Market Wise Policy Agenda," 1997. Inter-American Development Bank.

Chan, K. S., S. Mestelman, R. Moir, and R. A. Muller, "The Voluntary Provision of Public Goods under Varying Income Distributions," *Canadian Journal of Economics* 19 (1996):54–69.

Coase, Ronald H., "The Problem of Social Cost," *Journal of Law and Economics* 3 (October 1960):1–44.

Coleman, James S., "Free Riders and Zealots: The Role of Social Networks," *Sociological Theory* 6 (Spring 1988):52–57.

Craig, Ben and John Pencavel, "Participation and Productivity: A Comparison of Worker Cooperatives and Conventional Firms in the Plywood Industry," *Brookings Papers: Microeconomics* (1995):121–160.

Dandekar, V. M., "Crop Insurance in India: A Review 1976–l984-5," *Economic and Political Weekly* 20,25 (1985):A46–A59.

Dasgupta, Partha and Debraj Ray, "Inequality as a Determinant of Malnutrition and Unemployment: Theory," *Economic Journal* 96 (December 1986):1011–1034.

— and — , "Inequality as a Determinant of Malnutrition and Unemployment: Policy," *Economic Journal* 97 (March 1987):177–188.

Davis, Douglas D. and Charles A. Holt, *Experimental Economics* (Princeton: Princeton University Press, 1993).

Dayton-Johnson, J. and Pranab Bardhan, "Inequality and Conservation on the Local Commons: A Theoretical Exercise," 1997. Berkeley.

Deaton, Angus and Anne Case, "Large Cash Transfers to the Poor," 1997. Princeton University.

Demsetz, Harold and Kenneth Lehn, "The Structure of Corporate Control: Causes and Consequences," *Journal of Political Economy* 93,6 (1985):1155–1177.

Domar, Evsey and Richard A. Musgrave, "Proportional Income Taxation and Risk-Taking," *Quarterly Journal of Economics* 58 (1944):388–422.

Dong, Xioa-yuan and Gregory Dow, "Does Free Exit Reduce Shirking in Production Teams?," *Journal of Comparative Economics* 17 (1993):472–484.

— and — , "Monitoring Costs in Chinese Agricultural Teams," *Journal of Political Economy* 101,3 (1993):539–553.

Dutta, B., D. Ray, and K. Sengupta, "Contracts with Eviction in Infinitely Repeated Principal-Agent Relationships," in Pranab Bardhan (ed.) *The Economic Theory of Agrarian Institutions* (Oxford: Clarendon Press, 1989).

Elster, Jon, *The Cement of Society* (Cambridge: Cambridge University Press, 1989).

Engerman, Stanley L. and Kenneth L. Sokoloff, "Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A fiew from Economic Historians of the United States," December 1994. NBER Working Paper.

Evans, David and Boyan Jovanovic, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy* 97,4 (1989):808–827.

— and Linda Leighton, "Some Empirical Aspects of Entrepreneurship," *American Economic Review* 79,3 (June 1989):519–535.

Falk, Armin and Urs Fischbacher, "Kindness is the Parent of Kindness: Modeling Reciprocity," 1998. Institute for Empirical Economic Research, University of Zürich.

Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," 1997. Institute for Empirical Research in Economics, University of Zürich.

— and Simon Gächter, "Cooperation and Punishment," 1996. Working Paper, Institute for Empirical Economic Research, University of Zürich.

— , — , and Georg Kirchsteiger, "Reciprocity as a Contract Enforcement Device: Experimental Evidence," *Econometrica* 65,4 (July 1997):833–860.

Forsythe, Robert, Joel Horowitz, N. E. Savin, and Martin Sefton, "Replicability, Fairness and Pay in Experiments with Simple Bargaining Games," *Games and Economic Behavior* 6,3 (May 1994):347–369.

Frohlich, Norman, John Godard, Joe Oppenheimer, and Frederick Starke, "Employee vs. Conventionally Owned and controlled Firms: n Experimental Analysis," November 1997. University of Maryland.

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies* 60,1 (1993):35–52.

Gintis, Herbert, "Financial Markets and the Political Structure of the Enterprise," *Journal of Economic Behavior and Organization* 1 (1989):311–322.

Glaeser, Edward L. and DiPasquale, "Incentives and Social Capital: Are Homeowners Better Citizens?," *Journal of Urban Economics (*forthcoming).

Green, Leonard, Joel Myerson, David Lichtman, Suzanne Rosen, and Astrid Fry, "Temporal Discounting in Choice Between Delayed Rewards: The Role of Age and Income," *Psychology and Aging* 11,1 (1996):79–84.

Greenberg, Edward, *Workplace Democracy: The Political Effects of Participation* (Ithaca: Cornell University Press, 1986).

Grossman, Sanford J. and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94,4 (August 1986):691–719.

Güth, Werner and Peter Ockenfels, "Efficiency by Trust in Fairness? Multiperiod Ultimatum Bargaining Experiments with an Increasing Cake," *International Journal of Game Theory* 22,1 (1993):51–73.

Güth, Werner, R. Schmittberger, and B. Schwarz, "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization* 3 (May 1982):367–388.

Hansen, Daniel G., "Individual Responses to a Group Incentive," *Industrial Relations (*1997).

Hausman, Jerry, "Individual Discount Rates and the Purchase and Utilization of Energy-using Durables," *Bell Journal of Economics* 10,1 (Spring 1979):33–54.

Hayami, Yujiro and Jean-Philippe Platteau, "Resource Endowments and Agricultural Development: Africa vs Asia," June 1997. Faculté Universitaires Notre-Dame de la Paix.

Hoff, Karla, "Comment on "Political Economy of Alleviating Poverty: Theory and Institutions," by Timothy Besley," *Annual World Bank Conference on Development Economics (*1996):139–144.

— , "Market Failures and the Distribution of Wealth: A Perspective from the Economics of Information," *Politics & Society* 24,4 (December 1996):411–432.

— and Andrew B. Lyon, "Non-Leaky Buckets: Optimal Redistributive Taxation and Agency Costs," *Journal of Public Economics* 26 (1995):365–390.

Hoffman, Elizabeth, Kevin McCabe, and Vernon L. Smith, "Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology," April, 1996. Unpublished.

Hölmstrom, Bengt, "Moral Hazard in Teams," *Bell Journal of Economics* 7 (1982):324–340.

— and Jean Tirole, "The Theory of the Firm," in R. Schmalensee and R. Willig (eds.) *Handbook of Industrial Organization* (Amsterdam: North-Holland, 1988).

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen, "Entrepreneurial Decisions and Liquidity Constraints," *RAND Journal of Economics* 25,2 (Summer 1994):334–347.

— , — , and — , "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy* 102,1 (1994):53–75.

Homans, George, *Social Behavior: Its Elementary Forms* (New York: Harcourt Brace, 1961).

Hossain, M., "Credit for Alleviation of Rural Poverty: the Grameen Bank in Bangladesh," 1988. International Food Policy Research Institute Report 65.

Jarvis, Lovell, "The Unravelling of Chile's Agrarian Reform, 1973-1986," in William Thiesenhusen (ed.) *Searching for Agrarian Reform in Latin America* (Boston: Unwin–Hyman, 1989) pp. 240–265.

Johansen, Leif, "The Bargaining Society and the Inefficiency of Bargaining," *Kyklos* 32,3 (1979):497–522.

Johnson, D. G., "Resource Allocation under Share Contracts," *Journal of Political Economy* 58 (1950):110–23.

Kandel, Eugene and Edward P. Lazear, "Peer Pressure and Partnerships," *Journal of Political Economy* 100,4 (August 1992):801–817.

Keefer, P. and S. Knack, "Polarization, Property Rights and the Links between Inequality and Growth," October 1995. World Bank.

Kollock, Peter, "Transforming Social Dilemmas: Group Identity and Cooperation," in Peter Danielson (ed.) *Modeling Rational and Moral Agents* (Oxford: Oxford University Press, 1997).

Kotwal, A., "The Role of Consumption Credit in Agricultural Tenancy," *Journal of Development Economics* 18 (1985):273–96.

Kramer, Roderick and Marilynn Brewer, "Effects of Group Identity on Resource Use in a Simulated Commons Dilemma," *Journal of Personality and Social Psychology* 46,5 (1984):1044–1057.

Laffont, Jean Jacques and Mohamed Salah Matoussi, "Moral Hazard, Financial Constraints, and Share Cropping in El Oulja," *Review of Economic Studies* 62 (1995):381–399.

Lawler, Edward J. and Jeongkoo Yoon, "Committment in Exchange Relations: Test of A Theory of Relational Cohesion," *American Sociological Review* 61 (February 1996):89–108.

Lazear, Edward, "Performance Pay and Productivity," July 1996. NBER working paper #5672.

Legros, Patrick and Andrew F. Newman, "Wealth Effects, Distribution, and the Theory of Organization," *Journal of Economic Theory* (August 1996).

— and Steven A. Matthews, "Efficient and Nearly-Efficient Partnerships," *Review of Economic Studies* 68 (1993):599–611.

Leibenstein, Harvey, *Economic Backwardness and Economic Growth* (New York: Wiley, 1957).

Loury, Glen, "Intergenerational Transfers and the Distribution of Earnings," *Econometrica* 49 (1981):843–67.

Manning, Alan, "Imperfect Labour Markets, the Stock Market, and the Inefficiency of Capitalism," *Oxford Economic Papers* 44 (1992):257–271.

McAfee, R. Preston and John McMillan, "Optimal Contracts for Teams," *International Economic Review* 32,3 (August 1991):561–577.

Meyer, Jack, "Two-Moment Decision Models and Expected Utility," *American Economic Review* 77,3 (1987):421–430.

Moene, Karl Ove, "Poverty and Land Ownership," *American Economic Review* 81,1 (March 1992):52–64.

Molinas, José R., "The Impact of Inequality, Gender, Social Capital, and External Assistance on Local-level Collective Action," *World Development (*1998).

Mookherjee, Dilip, "Informational Rents and Property Rights in Land," in John Roemer (ed.) *Property Rights, Incentives, and Welfare* (London: MacMillan, 1997).

Nerlove, Marc and Tjeppy D. Soedjiana, "Slamerans and Sheep: Savings and Small Ruminants in Semi-Subsistence Agriculture in Indonesia," 1996. Department of Agriculture and Resource Economics, University of Maryland.

Newbery, David, "Agricultural Institutions for Insurance and Stabilization," in Pranab Bardhan (ed.) *The Economic Theory of Agrarian Institutions* (Oxford: Oxford University Press, 1989).

North, Douglass C., *Institutions, Institutional Change and Economic Performance* (Cambridge: Cambridge University Press, 1990).

Olson, Mancur, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, MA: Harvard University Press, 1965).

Ostrom, Elinor, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge: Cambridge University Press, 1990).

— , James Walker, and Roy Gardner, "Covenants with and without a Sword: Self-Governance is Possible," *American Political Science Review* 86,2 (June 1992):404–417.

Perotti, Roberto, "Fiscal Policy, Income Distribution, and Growth," November 1992. Columbia Working Paper 636.

— , "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies* 60 (1993):755–66.

Piketty, Thomas, "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies* 64,2 (April 1997):173–189.

Platteau, Jean-Philippe, J. Murickan, A. Palatty, and E. Delbar, "Rural Credit Market in a Backward Area: A Kerala Fishing Village," *Economic and Political Weekly (*October 1980):1765–80.

Prosterman, Roy L. and Jeffrey M. Riedinger, *Land Reform and Democratic Development* (Baltimore, MD: Johns Hopkins University Press, 1979).

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 83,5 (1993):1281–1302.

Richards, John F., *The Mughal Empire* (Cambridge: Cambridge University Press, 1993).

Robinson, James, "Distribution and Institutional Structure: Some Preliminary Notes," 1996. unpublished.

Rosenzweig, Mark and Hans P. Binswanger, "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *Economic Journal* 103 (January 1993):56–78.

Rosenzweig, Mark R. and Kenneth I. Wolpin, "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investment in Bullocks in India," *Journal of Political Economy* 101,2 (1993):223–244.

Rotemburg, Julio J., "Human Relations in the Workplace," *Journal of Political Economy* 102,4 (1994):684–717.

Runge, C. F., "Common Property Externalities: Isolation, Assurance, and Resource Depletion in a Traditional Grazing Context," *American Journal of Agricultural Economics* 63 (1981):595–606.

Saha, Atanu, Richard C. Shumway, and Hovav Talpaz, "Joint Estimation of Risk Preference Structure and Technology Using Expo-Power Utility," *American Journal of Agricultural Economics* 76,2 (May 1994):173–184.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls, "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy," *Science* 277 (August 15, 1997):918–924.

Sanderatne, Nimal, "The Political Economy of Small Farmer Loan Delinquency," *Savings and Development* 10 (1986):343–353.

Sappington, David, "Limited Liability Contracts between Principal and Agent," *Journal of Economic Theory* 29,1 (February 1983):1–21.

Shaban, Radwan Ali, "Testing Between Competing Models of Sharecropping," *Journal of Political Economy* 95,5 (1987):893–920.

Shelton, John, "Allocative Efficiency vs X-Efficiency: Comment," *American Economic Review* 57,5 (1957):1252–1258.

Shetty, S., "Limited Liability, Wealth Differences, and the Tenancy Ladder in Agrarian Economies," *Journal of Development Economics* 29 (1988):1–22.

Shiller, Robert J., *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks* (Oxford: Clarendon Press, 1993).

Sial, Maqbool and Michael Carter, "Financial Market Efficiency in an Agrarian Economy: Microeconometric Analysis of the Pakistani Punjab," *Journal of Development Studies* 32,2 (June 1996):771–798.

Sinn, Hans-Werner, "Expected Utility, μ-σ Preferences, and Linear Distribution Classes: A further Result," *Journal of Risk and Uncertainty* 3 (1990):277–281.

— , "A Theory of the Welfare State," *Scandinavian Journal of Economics* 95,4 (1995):495–526.

Stiglitz, Joseph, "Incentives and Risk Sharing in Sharecropping," *Review of Economic Studies* 41,2 (April 1974):219–255.

— , "Rational Peasants, Efficient Institutions, and a Theory or Rural Organization," in Pranab Bardhan (ed.) *The Economic Theory of Agrarian Institutions* (Oxford: Oxford University Press, 1989) pp. 10–29.

Stiglitz, Joseph E., "Peer Monitoring and Credit Markets," in Karla Hoff, Avishay Braverman, and Joseph E. Stiglitz (eds.) *The Economics of Rural Organization: Theory, Practice, and Policy* (New York: Oxford University Press, 1993) pp. 70–85.

Svensson, J., "Investment, Property Rights and Political Instability : Theory and Evidence," 1993. Stockolm University, Institute for International Economic Studies.

Tilly, Charles, "Charivaris, Repertoires and Urban Politics," in John M. Merriman (ed.) *French Cities in the Nineteenth Century* (New York: Homs and Meier, 1981) pp. –.

Varian, Hal R., "Monitoring Agents with Other Agents," *Journal of Institutional and Theoretical Economics* 46,1 (1990):153–174.

Venieris, Y. and D. Gupta, "Income Distribution and Sociopolitical Instability as Determinants of Savings: A Cross-Sectional Model," *Journal of Political Economy* 94 (1986):873–883.

Verba, Sidney, Kay Lehman Schlozman, and Henry Brady, *Voice and Equality: Civic Voluntarism in American Politics* (Cambridge, MA: Harvard University Press, 1995).

Weissing, Franz and Elinor Ostrom, "Irrigation Institutions and the Games Irrigators Play: Rule Enforcement without Guards," in Reinhard Selten (ed.) *Game Equilibrium Models II: Methods Morals and Markets* (Berlin: Springer-Verlag, 1991) pp. 188–262.

Whyte, William F., *Money and Motivation* (New York: Harper & Row, 1955).

Williams, Jeffrey R., Gordon L. Carriker, G. Art Barnaby, and Jayson K. Harper, "Crop Insurance and Disaster Assistance Designs for Wheat and Grain Sorghum," *American Journal of Agricultural Economics* 75 (May 1993):435–447.

Zimmerman, Frederic and Michael Carter, "Dynamic Portfolio Management in Developing Countries," May 1997. University of Wisconsin.

e:\papers\handbook\Wealth Constraints and Economic Performance.tex